Utilitas (2026), **38**, 289–302
doi:10.1017/S0953820826100429

CAMBRIDGE UNIVERSITY PRESS

ARTICLES

Enduring Objections to Actual Rule Consequentialism

Vuko Andrić

Linköping University, Sweden and Institute for Futures Studies, Sweden
Email: vuko.andric@liu.se

(Received 8 July 2025 UTC; revised 26 March 2026 UTC; accepted 23 May 2026 UTC; first published online 24 July 2026 UTC)

Abstract

According to rule consequentialism, the moral status of actions is determined by their conformity to authoritative moral rules, and the authoritativeness of rules is determined by their consequences. While ideal rule consequentialism focuses on the consequences the adoption of rules would have, actual rule consequentialism focuses on occurrent consequences. In this paper, I criticize two recent versions of actual rule consequentialism. According to Richard B. Miller's account, actions are morally right unless they are proscribed by sufficiently weighty, legitimate ethical rules that are actually adopted in the agent's society. According to Caleb Perl's Patterned View, while rules can be authoritative without being actually adopted, the evaluation of rules turns on the occurrent consequences of actually obtaining behavior. I argue that both Miller's and Perl's views share the problems of earlier versions of actual rule consequentialism, originally identified by D.H. Hodgson and Fred Feldman: they are overly conservative and objectionably conventionalist.

Keywords: rule consequentialism; actual rule consequentialism; patterned view

According to rule consequentialism, the moral status of actions is determined by their conformity to authoritative moral rules, and the authoritativeness of rules is determined by their consequences. The distinction between ideal and actual rule consequentialism concerns how rules are assessed. According to the predominant version, ideal rule consequentialism, the rules whose adoption would have the (prospectively) best consequences are authoritative. Actual rule consequentialism, by contrast, holds that rules are authoritative in virtue of good consequences that actually occur.[1]

[1]I here deliberately speak of "good consequences" rather than "good consequences of rules" because I think that Perl's (2021) view, which focuses on the occurrent consequences of actual patterns of behavior, should be classified as a version of actual rule consequentialism. More on this below.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

Focusing on the consequences of actually adopted rules – rather than rules that merely could be, but possibly are not, adopted – promises certain advantages. First, a possible motivation for accepting actual rule consequentialism arises from a combination of generic consequentialism – understood as the claim that the point of morality is to promote good consequences – and dissatisfaction with both act consequentialism and ideal rule consequentialism. The implications of act consequentialism are often considered implausible.[2] Ideal rule consequentialism faces well-known objections, most notably those arising when the rules it would be ideal to adopt are not, in fact, adopted in society.[3] Actual rule consequentialism promises to avoid both kinds of problem (Johnson 1991; Miller 2009; Perl 2021).

Second, our moral intuitions are shaped by the rules actually adopted in society. Actual rule consequentialism may score points on a reflective equilibrium account of moral theory selection by accommodating these intuitions.[4]

Third, there are instructive analyses demonstrating the good consequences of actually adopted rules.[5] These analyses support a claim Mill makes in arguing that people should not try to apply the utility principle directly in everyday life, but should instead rely on rules that come from historical experience:

> During all that time mankind have been learning by experience the tendencies of actions; on which experience all the prudence, as well as all the morality of life, is dependent. [. . .] mankind must by this time have acquired positive beliefs as to the effects of some actions on their happiness; and the beliefs which have thus come down are the rules of morality for the multitude, and for the philosopher until he has succeeded in finding better.[6]

Actual rule consequentialism may be better positioned to account for such findings than either act consequentialism or ideal rule consequentialism.[7]

However, actual rule consequentialism faces its own problems, which stem from the fact that the rules actually adopted in society may fall far short of the ideal. Two recent versions – those defended by Richard B. Miller (2009) and Caleb Perl (2021) – attempt to address these problems. In this paper, I examine their contributions. My claim is that, ultimately, these recent versions of actual rule consequentialism, despite the careful and innovative arguments advanced in their favor, fall prey to the same kinds of problems that afflict more traditional formulations. In particular, I argue that the newer versions remain overly conservative and objectionably conventionalist.

The paper is structured as follows. Section 1 outlines the historical background and introduces the early criticisms of actual rule consequentialism developed by D.H. Hodgson (1967) and Fred Feldman (1978). Section 2 argues that Miller's version of rule

[2]Textbooks offering instructive overviews include Bykvist (2009, chs. 5–9); Mulgan (2007, ch. 5); and Shaw (1999, ch. 4).

[3]For an overview of objections to ideal rule consequentialism, see Hooker (2023).

[4]On the method of reflective equilibrium, see Knight (2025), on applying the method to justify ideal rule consequentialism, see Hooker (1996, 2000, ch. 1).

[5]Examples of the analyses I have in mind include Hardin (1988); Bailey (1997); and Johnson (1991).

[6]Mill ([1861] 1998, ch. 2, par. 24). Compare Miller (2009), whose elaborations on moral theory include frequent appeals to Mill.

[7]Of course, act and ideal rule consequentialists should also accommodate such analyses, which is a complex task as instructively demonstrated in Shaw (1999).

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

consequentialism does not overcome the problems identified by Hodgson and Feldman. Section 3 contends that the same holds for Perl's view. Section 4 concludes.

1 Background

Predecessors of actual rule consequentialism, which emphasized the occurrent good consequences of actually adopted rules, were defended by Harrod (1936) and Rawls (1955), and criticized by Hodgson (1967) and Feldman (1978).

Hodgson (1967, pp. 26–32, esp. 29–30) emphasizes two problems. First, when the conventions actually adopted in a society are suboptimal, actual rule consequentialism risks being overly conservative, as it does not require efforts to actively improve society's conventions. Second, when actual conventions are uncertain or conflicting, actual rule consequentialism appears incomplete and risks collapsing into either act consequentialism or ideal rule consequentialism.

Feldman (1978, pp. 71–77) argues that Rawls's proposal collapses into either ideal rule consequentialism or conventionalism (the view that any rules actually adopted by a society are authoritative for its members), and maintains that conventionalism is implausible, since conventions can be morally abhorrent – as exemplified by slavery, torture, human sacrifice, and cannibalism.

A fully developed version of actual rule consequentialism, which seeks to avoid some of these problems, was defended by Conrad Johnson (1991). A distinctive feature of Johnson's theory is the *primary principle*, which has an act-consequentialist character in that it requires agents "to promote good and avoid evil." The primary principle is intended to fill the gaps where actually adopted rules fail to evaluate certain kinds of actions. It also provides the baseline for assessing the justification of actually adopted rules. A rule is justified, Johnson (1991, p. 38) argues, if its general acceptance produces greater overall good than would be produced if actions in that domain were instead governed by the primary principle.

Miller (2009) offers a striking critique of Johnson's rule consequentialism, focusing on Johnson's primary principle. Since the primary principle is meant to apply when actually adopted rules leave gaps, Johnson's theory inherits problems typically associated with act consequentialism. As Miller (2009, p. 14) observes,

> a reasonable ethics will forbid certain classes of actions while at the same time creating a generous sphere of permissible actions in which agents can do as they please. The gaps between the rules are spaces in which responsible persons are free. [. . .] Johnson should note that the security from blame which he prizes is lost if his primary principle is allowed to function as a rule.

Miller is right: if actual rule consequentialists wish to avoid concerns related to the demandingness objection against act consequentialism, they should not use the primary principle to fill such gaps.

Insofar as Johnson's theory evaluates actually adopted rules by comparing their consequences to those that would result if people followed the primary principle, it appears unable to fully account for the motivational limitations of human agents. As Miller (2009, p. 15) notes, "we must realistically predict what normally conscientious people would do in the absence of the rule. Realistically we ought to expect that in the situation formerly covered by the rule individuals would not revert to act [consequentialism] but rather would

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

pursue self-interest or the interests of friends and family."[8] I agree with Miller that rule consequentialism must consider the motivational limitations of human agents when evaluating moral rules, and that conformity with the primary principle does not provide a plausible baseline for such evaluations.

Miller's critique of Johnson's theory is, in this respect, convincing. But does Miller's own version of actual rule consequentialism fare any better?

2 Miller's actual rule consequentialism

Miller (2009, p. 22) proposes the following criterion of rightness for his version of actual rule consequentialism: "An act is morally right either if it is not proscribed by any legitimate ethical rule in the society in which it would occur, or if so proscribed it is also required by another legitimate ethical rule of equal or greater weight." While it is unclear to me why Miller formulates only a sufficient condition ("if") rather than a necessary and sufficient condition ("iff") for moral rightness, I set this issue aside for the present purposes. Two key expressions Miller uses – "legitimate ethical rule" and "equal or greater weight" – require further clarification.

According to Miller (2009, p. 22), a legitimate ethical rule is "an ethical practice that is de facto in a given society, contains a disaster avoidance qualification, and promotes the happiness of people affected by the rule." Furthermore, Miller explains that, to promote happiness, a rule "has to make life better than it would be otherwise." But how would it be otherwise? The relevant baseline, according to Miller (2009, p. 15), is given by what he calls the realistic assumption that, in the absence of the rule, individuals "would pursue self-interest or the interests of friends and family."

Miller (2009, p. 22) defines that a rule has greater weight (relative to a society) than another "if both rules are de facto in a given society and if mature individuals in the society generally agree that the first rule takes precedence over the second." Two rules have equal weight "if mature individuals in the society do not generally agree that one of the rules takes precedence over the other." Applying these definitions to Miller's criterion of moral permissibility reveals that his theory is quite liberal, offering a "generous sphere of permissible actions": an act is morally right if it is not proscribed by a legitimate ethical rule, and it is also morally right if it is proscribed by one but conflicts with another legitimate ethical rule where mature individuals do not generally agree that the first rule takes precedence over the second.

Let us now turn to an assessment of Miller's proposal. Miller repeatedly emphasizes the beneficial consequences of not being subject to moral obligations (Miller 2009, pp. 19, 21, 23, 24). Accordingly, unlike Johnson's, Miller's version of actual rule consequentialism does not fill gaps in actually adopted rules by introducing a moral principle akin to act consequentialism, thereby avoiding problems associated with that doctrine. Moreover, given the realistic baseline discussed earlier, Miller's rule consequentialism appears to avoid the utopianism or ideal world problem: the difficulty faced by ideal rule consequentialism whereby, if too few people actually adopt the authoritative moral rules, compliance with these rules can often seem morally wrong.[9]

However, I find Miller's rule consequentialism too conservative and too close to conventionalism to be fully convincing. To illustrate this, I will present three cases.

[8]I have replaced "utilitarianism" with "consequentialism."

[9]It is debatable if Miller's rule consequentialism really avoids the problem: see Podgorski (2018, p. 291).

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

Harsh Promising

> In Society, the actually adopted rules of promising are overly strict: promisors are bound to keep their promises even when doing so imposes very high costs on them. Life under these harsh promising rules is better than it would be if there were no practice of promising at all, and individuals merely pursued self-interest or the interests of friends and family. Nevertheless, adopting more permissive promising rules – similar to those endorsed by our common-sense morality, which allow promisors greater leeway – would yield better consequences for Society. Furthermore, the members of Society could relatively easily be persuaded to adopt these superior promising rules.

Actual rule consequentialists express skepticism about unrealistic cases (Miller 2009, p. 27; Perl 2023). Although I do not share this skepticism, this is not the place to debate it. Instead, I will emphasize that the cases I present are not unrealistic. This is readily confirmed in the case of *Harsh Promising*. Different individuals (including philosophers – Kant being a notable example) and different societies (across various times) genuinely disagree about how high the costs to the promisor must be for them to be "off the hook."[10]

Miller's rule consequentialism implies that Society's promising rules are both legitimate and suboptimal. They are legitimate because life under these harsh rules is better than it would be in their absence – i.e., if there were no practice of promising and individuals simply pursued self-interest or the interests of friends and family. As such, the rules are binding in Society according to Miller's framework. However, they remain suboptimal, since the adoption of more permissive rules – such as those endorsed by our common-sense morality, which allow promisors greater leeway – would yield better consequences.

What should rule consequentialists say about *Harsh Promising*? I take it to be a case in which common-sense (pre-theoretical) intuitions and generic consequentialist reasoning align. It seems plausible that members of Society – not only collectively, but also individually – ought to endorse and adopt the better rules about promising, namely the more permissive rules we ourselves follow, thereby helping to promote a transition.

Naturally, there are caveats. Members of Society should communicate this shift to those with whom they interact in order to avoid frustrating expectations. Moreover, since promises made prior to the shift were made under the assumption that the harsh rules are binding, promisors may have moral duties to keep some or all of them, especially in high-stakes cases or where breaking the promise would be unfair to the promisee.

Ideal rule consequentialism seems to align with, and even to explain, our common-sense moral intuitions in Harsh Promising. If the permissive rules about promising are optimal, the members of Society should adopt them. Ideal rule consequentialism also seems to have the resources to explain the caveats just mentioned. As Hooker (2000, p. 122) puts it: "Just as in a world where things can go badly wrong we need a strong 'prevent disaster' rule, so we need a rule requiring us to comply with conventional suboptimal rules when our not doing so would cause serious unfairness." This accounts for high-stakes cases and unfairness toward promisees.[11]

[10]Empirical work on different views of promises includes Ma et al. (2020) and Friedman et al. (2018).

[11]Chapter 5 of Hooker (2000) is devoted entirely to the role of conventions in ideal rule consequentialism.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

The conclusion that members of Society ought to accept and comply with the optimal rules is also supported by a generic consequentialist perspective. If the point of morality is ultimately to make things go best, this further suggests that members of Society should adopt the better rules about promising and thereby contribute to improving their society. (Again, with possible exceptions, which would not be governed by rules about promising but by other rules, such as the "prevent disaster" rule.)

That said, appeals to generic consequentialism must be made with care. Hooker (1996; 2000: ch. 1) famously rejects such appeals when assessing rule consequentialist theories. While other rule consequentialists – most notably Harsanyi (1977a, 1977b, 1993) – disagree with Hooker, I cannot simply presuppose generic consequentialism in order to critique Miller's version of rule consequentialism.

Here, then, is how I see the dialectical situation. I first noted that ordinary moral intuitions support the verdict that members of Society should adopt more permissive, conventional rules of promising – rules like our own. Generic consequentialist considerations serve an auxiliary role: for those who endorse generic consequentialism, it matters that such reasoning also supports this verdict. Furthermore, this means that Miller cannot appeal to generic consequentialism to defend the prima facie implausible implications his theory yields in the *Harsh Promising* case.

What, then, can Miller appeal to in order to defend his theory's implications? If the case cannot be dismissed as unrealistic, and if both substantive moral intuitions and generic consequentialism are off the table, then only intuitions supporting a form of conventionalism appear to remain. The suggestion would have to be that the harsh promising rules are binding in Society simply because they are actually adopted there.[12]

But this suggestion – that the harsh promising rules are binding simply because they are actually adopted – would require a strong kind of conventionalism. Recall the quote by Mill, provided in the introduction, according to which humankind, in their history, "have acquired positive beliefs as to the effects of some actions on their happiness; and the beliefs which have thus come down are the rules of morality for the multitude, and for the philosopher until he has succeeded in finding better." Many utilitarians might agree with this claim if it is understood as pointing out the *heuristic/epistemic* relevance of actually adopted rules regarding the promotion of well-being. In cases like *Harsh Promising*, however, there is no such relevance and facts about actually adopted rules would instead need to be understood as *normative grounds*, independent of good consequences, to support Miller's position.

All things considered, for those of us who do not share strong conventionalist intuitions, Miller's version of actual rule consequentialism does not seem to deliver the right

[12]A reviewer helpfully pointed out that Miller (2009, p. 7) subscribes to "satisficing" – rules need only be "good enough" – which suggests that my arguments may be question-begging, since I claim that the authoritative rules should be optimal rather than merely good enough. However, while an endorsement of satisficing could perhaps be justified by appeal to moral intuitions or on consequentialist grounds (Slote and Pettit 1984), accepting satisficing as a fundamental principle, as would be required here, is implausible. A further objection concerns my methodology, since Miller (2009, p. 7) holds that "the criterion for evaluating moral theory is revisionary and pragmatic and not descriptive or based on intuitions." While a full discussion of Miller's revisionary criterion lies beyond the scope of this paper, two brief points are worth noting. First, it is doubtful that Miller can justify central tenets of his theory – the value of well-being, the distinction between legitimate and illegitimate rules, and the satisficing approach – without appeal to moral intuitions. Second, it is hard to see how a pragmatic approach could justify more than two tenets on which utilitarians can agree anyway: the importance of well-being and the heuristic or epistemic value of actually adopted rules.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

verdict in *Harsh Promising*. In that case, it appears more plausible that the optimal rules – rather than those merely adopted by Society – should be regarded as authoritative.

The second case is, again, a thought experiment:

> *Religious Dystopia*
>
> The majority of people in our society convert to a new religion. As a result, many central rules of common-sense morality are no longer widely adopted. Key domains such as healthcare, the economy, and child education are now largely managed by a priestly class in ways that diverge sharply from previously accepted special obligations. Healthcare no longer involves individual care; instead, it consists solely of priests offering prayers for the health of believers. Children are separated from their parents and raised collectively by the priests. Factories are operated with the exclusive goal of maximizing short-term profit, with little regard for worker well-being.
>
> The remaining group of non-believers is radically dispersed and socially isolated; there are no sub-communities of dissenters. Some privately owned factories still exist, along with occasional instances of parental child-rearing and a handful of medical doctors still treating patients. These isolated practices are tolerated by the religious authorities but remain marginal.

The case seems realistic enough. Admittedly, it is unlikely that contemporary societies would convert to the religion described in *Religious Dystopia*. However, history offers no shortage of religious and political ideologies that many of us now regard as extreme or irrational. We are familiar with the kinds of ideas that have given rise to institutions resembling those described in *Religious Dystopia*.[13]

If we assume (plausibly) that the predominant societal arrangements in *Religious Dystopia* are illegitimate in Miller's sense, then his rule consequentialism avoids false positives. The problem, however, is that it also implies that the non-believers in *Religious Dystopia* do not have traditional special obligations. This is because the rules that ground such obligations are no longer actually adopted. Of course, Miller's theory would generate the desired special obligations if the scattered instances of traditional (non-religious) childcare, healthcare, and economic practices were to form sub-societies. But by hypothesis, that is not the case: there are simply too few instances of such traditional practices to constitute a rule-governed social structure.[14]

Miller's rule consequentialism fails by implying that traditional special obligations do not exist in *Religious Dystopia* – a false negative. Common-sense morality suggests that special obligations persist in the isolated cases of parental care, privately owned factories, and practicing doctors. We can plausibly assume that if these remaining factory owners, parents, and doctors adopted rules embodying the special obligations recognized by

[13]Readers who still find *Religious Dystopia* too fanciful can easily construct a more realistic case by focusing on just one of the described changes: in healthcare, the economy, or child education.

[14]Miller's rule consequentialism does not merely fail to imply that the non-believers have the relevant special obligations; it positively implies that they do not. Recall that, according to Miller (2009, p. 22), an act is morally right if it is not proscribed by any legitimate rule actually adopted in the society where it occurs. If the rules constituting the relevant special obligations are not adopted in that society, then those rules cannot render acts violating these obligations morally wrong. Since no other rules prohibiting such acts are adopted either, it follows that these acts are morally right.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

common-sense morality, the consequences would be better. Thus, both substantive moral intuitions and generic consequentialism speak against Miller's theory. It is overly conservative – too conservative for me at any rate.[15]

The third case challenging Miller's rule consequentialism concerns slavery. Miller (2009, pp. 24–25) contends that, unlike conventionalism, his theory does not imply that slaves have special obligations to their masters. Recall that the baseline for assessing the (il)legitimacy of actually adopted rules is the assumption that, without such rules, individuals would pursue self-interest or the interests of friends and family. But does Miller's theory truly avoid the conclusion that slaves owe special obligations to their masters? This depends on the circumstances. The problem here is well-known from contractarian discussions (Gauthier 1986, p. 195).

Slavery

An ethics of slavery is adopted in Society, which consists of two classes: rulers and slaves. The rulers control coercive means, such as weapons, and occasional disobedience by individual slaves is punished. However, since both rulers and slaves accept that slaves have moral obligations to obey their masters, enforcement and punishment are rarely necessary. If the rules of slavery were not adopted and people instead pursued self-interest and the interests of friends and family, significantly more coercion and punishment would be required to maintain slavery.

What this comparison shows is that adopting a slave ethics can be a Pareto improvement over a situation where slavery is violently enforced because slaves do not believe they have a moral obligation to obey their masters. Coercion and punishment impose costs on both masters and slaves. Consequently, in some circumstances Miller's theory cannot avoid the conclusion that slaves ought to obey their masters. While a defender of Miller's actual rule consequentialism might be willing to accept this implication – similar to some contractarians – for most of us, I think, this conclusion is not acceptable.

In sum, all three cases highlight significant problems for Miller's actual rule consequentialism, suggesting that it is overly conservative and endorses unjustified conventions. This aligns with the criticisms Hodgson and Feldman raised against earlier versions of actual rule consequentialism. In *Harsh Promising*, Miller's theory is too conservative because it fails to require individuals to actively improve society's conventions. In *Religious Dystopia*, it is too conservative by not demanding the adoption of better rules. Finally, Miller's rule consequentialism is overly conventionalist in *Slavery*, as it legitimizes an actually adopted ethics of slavery.

3 Perl's patterned view

Unlike the views defended by Johnson and Miller, Perl's rule consequentialism, which he calls "The Patterned View," does not focus on actually adopted rules, but rather on

[15]To be fair, this won't be a surprise for Miller (2009) himself, who writes on p. 28: "I believe that my understanding of the nature and purpose of moral theory is more consistent with the great and ancient tradition of philosophical thought from Socrates to Mill. Since Mill's time the audience for philosophy has gradually narrowed and philosophy has become mostly a private conversation among the cognoscenti. I retain the small and probably naïve hope that a more pragmatic approach to philosophical method might win some of that audience back."

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

actual patterns of behavior. Nonetheless, Perl's rule consequentialism is related to those defended by Johnson and Miller and qualifies as a version of actual rule consequentialism, understood as the position that rules are authoritative in virtue of occurrent good consequences. The key difference is that, according to Perl, the occurrent good consequences need not arise from actually adopted rules but from actually occurring patterns of behavior.

According to Perl (2021, p. 98), "the set of morally best rules is the set of sufficiently general rules that has optimal moral value." Perl's evaluation of moral rules involves two key requirements: the Generality Requirement, which states that the morally best rules must be sufficiently general, and the Consilience Requirement, which holds that the moral value of a rule *R* is determined by everything good or bad that is caused by actions that *R* classifies as morally right.

To illustrate the Patterned View, Perl (2021, p. 92) uses a well-known example: "You are a sheriff in a small southern town. A murder has been committed, and you do not have the least idea who committed it, but a lynch mob will hang five others if you do not fasten the crime to one individual." How does Perl's view explain our judgment that it is wrong for me to scapegoat someone, even if doing so would have the best consequences? Given the Generality Requirement, the authoritative rules will have a general form:

> *The Scapegoating Rule*: scapegoating is right when the odds of discovery are low enough, and fair trials are wrong.
>
> *The Fair Trial Rule*: scapegoating is wrong even when the odds of discovery are low enough, and fair trials are always right.

Here, then, is what the Patterned View implies for the case:

> Given the Consilience Requirement, the great good of a trusted justice system is credited to the fair trial rule and is not credited to the scapegoating rule. Fair trials are what sustain trust in the justice system. But according to the scapegoating rule, fair trials are wrong when the person could be scapegoated. So the actions that cause trust in the justice system are wrong according to the scapegoating rule – which means that the great good of a trusted justice system isn't credited to the scapegoating rule. (Perl 2021, pp. 99–100).

Of course, the survival of the five people is attributed to the scapegoating rule rather than the fair trial rule. However, since we assume that the overall good produced by a trusted justice system outweighs all the benefits gained from successful scapegoating, Perl's view ultimately vindicates the fair trial rule.

Is the Patterned View convincing? Perl's theory is more flexible than those of Johnson and Miller because it allows not only actually adopted rules but also rules relevantly correlated to actual patterns of behavior to be authoritative. This flexibility enables Perl's theory to better address cases like *Harsh Promising*, *Religious Dystopia*, and *Slavery*. Nevertheless, I will argue that Perl's theory remains too conservative.[16]

My critique is based on a challenge that Perl himself (2021, pp. 123–124) attempts to rebut. The challenge compares an existing patronage system for jobs with a merely

[16]At the same time, Perl's theory might also be too utopian in some respects: see Rumbold (2024).

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

possible merit-based system: "The extant patronage system produces actual good, while the merely possible system doesn't, because it's not in place. But the merit-based system plausibly would produce a great deal more good. And the Consilience Requirement credits rules only with actual good." The apparent implication is that there is no requirement to switch to the merit-based system. The challenge for Perl, then, is to demonstrate that the Patterned View indeed sanctions such a switch.

Perl offers three responses, all of which I find unconvincing. His first response (2021, p. 123) claims that "the challenge is only forceful if society never switches to the better system. If it does switch at some point, the Consilience Requirement will credit rules enjoining the switch with the future actual good, and thus it predicates a requirement to switch." Perl here seems to suggest that the consequences of all actions – past, present, and future – are attributed to rules classifying those actions as right. However, this reading of the Consilience Requirement is problematic.[17] Since circumstances change over time, the past and future performances of some type of action might have excellent consequences while currently performing actions of that type would have bad consequences. As a result, a rule endorsing this type of action for the present could implausibly be deemed authoritative under the Patterned View.

To illustrate the problem, a few examples may help. The practice of senicide might have been optimal in the past under certain harsh conditions, but it is clearly a bad idea under present circumstances. Allocating organs strictly according to need, without requiring donor consent, might be justified in a future society that has achieved significant bio-enhancement – such that people are far more altruistic – but certainly not before. Likewise, socialism might become an optimal social arrangement after sufficient technological progress has reduced the importance of profit motives and competition for innovation and productivity, but it could have poor consequences in earlier stages of development. These examples highlight the difficulty facing a temporally neutral version of the Patterned View, one in which the consequences of all past, present, and future actions are credited to rules. Such a view risks authorizing rules whose endorsement is clearly inappropriate in the present.

In any case, even if we grant that the consequences of all actions ever performed – past, present, and future – are to be ascribed to rules that classify those actions as right, the challenge remains. As Perl himself acknowledges, we can simply assume that society never switches from the patronage system to a merit-based one.

Perl's second response to the challenge proceeds as follows:

> Crucially, though, the patronage system still violates the Benefit Rule: the patronage system involves individuals choosing a lesser benefit over a greater. So the Patterned View still predicts that choosing based on merit is what's right. But the Patterned View does not give a distinctively patterned explanation of the wrong of the patronage system, since the Benefit Rule doesn't reflect goods distinctive to patterns. (Perl 2021, p. 123)

Unfortunately, this response calls for further elaboration. What exactly is the Benefit Rule, and how is it related to other rules in the theory? I will quote the relevant passage from Perl (2021), as doing so will make it clear why this second response also fails.

[17]A problem I won't press here is that future actions on some views cannot be assigned to rules because the future is open in that there is, at present, no fact about which actions will be performed.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

When considering the possibility that individuals sometimes achieve more good on their own than would be achieved by collective patterns of behavior, Perl (2021, p. 122) explains:

> Maybe in some areas, individuals striking out is what's best – some obligations of beneficence may be correctly explained in that way – so let's assume that individuals striking out is sometimes best. Given that assumption, the Patterned View will also incorporate an act-consequentialist-style rule:
>
> > Benefit Rule: choosing a greater benefit over a lesser benefit is right.
>
> Crucially, though, the Patterned View will subordinate the Benefit Rule to the other rules we've discussed.
>
> > Fair Trials Rule: scapegoating is wrong even when the odds of discovery are low enough, and fair trials are always right.
>
> It'll subordinate the Benefit Rule by taking it to only apply to actions that already conform to the other rules.

Let us now bring the pieces together. The Benefit Rule functions as an act-consequentialist-style rule. This explains Perl's second response to the challenge: the patronage system violates the Benefit Rule because it involves individuals choosing a lesser benefit over a greater one.

However, it is important to understand how the Benefit Rule relates to the other authoritative rules in Perl's framework. His elaborations make clear that the Benefit Rule is subordinate to the general rules identified through the Patterned View. Perl is committed to this hierarchy because, otherwise, the Benefit Rule would justify scapegoating – an implication he himself rightly rejects.

Yet this subordination undermines Perl's second response to the patronage system challenge. Since the Benefit Rule is subordinate to the other authoritative rules, it applies only to actions that conform to those rules – in this case, the rules of the patronage system. The Benefit Rule might help Perl address the challenge if patronage appointments were merely permissible under the system, but not if they are required. If the system requires patronage appointments, then the Benefit Rule becomes inapplicable, given its subordinated status. Of course, there is no reason to assume that the patronage system merely permits rather than requires patronage appointments. To evaluate the challenge meaningfully, we must consider it in its strongest form. Consequently, Perl's second response fails as well.[18]

Finally, here is Perl's third response:

> I've been granting that the patronage system will last until the end of time. But we wouldn't know one way or another; the Patterned View needs fleshing out to guide ignorant creatures like us. Some versions of the Patterned View could ground our

[18]A reviewer helpfully pointed out that, for the Patterned View to sanction the patronage system, it must also be assumed that the patronage system achieves some good different from that achieved by people following the Benefit Rule. Realistic examples include goods related to loyalty, predictability, and motivation. We should assume that this is so, again in order to consider the challenge in its strongest form.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

> moral obligations in patterns with a reasonable chance of being actual. Such versions would classify patronage as wrong so long as there's a reasonable chance that a merit-based alternative comes to exist. Those versions go even further in eliminating status quo bias. (Perl's 2021, p. 124)

In the introduction, I mentioned several possible motivations for pursuing actual rule consequentialism. One major motivation, however, which I have not yet discussed, is emphasized throughout Perl's (2021) paper. Perl's stated aim is to develop a version of rule consequentialism that avoids "counterfactual comparisons about how things could go differently" (p. 89). Yet it is unclear how this aim is compatible with a theory that evaluates "patterns with a reasonable chance of being actual" (p. 124).

Such a notion risks reintroducing counterfactual reasoning in a disguised form, since determining whether a pattern has a reasonable chance of becoming actual might rely on thinking about how things *could* go differently. At least, this would seem to be the case if the occurrence of some pattern in the future depends on what we decide to do in the present. There might be a very high probability (a "reasonable chance") of a merit-based system coming to exist *conditional on* sufficiently many people now adopting the rules associated with the system, but a very low probability of the merit-based system coming to exist *conditional on* sufficiently few people now adopting these rules. If *we* are the people in question, it is controversial whether we can move from conditional to unconditional probabilities (cf. Rabinowicz 2002). But even if we can, our thereby acquired reflexive third-person perspective (as observers of ourselves) might well piggyback on our first-person perspectives (as practical deliberators) on the question of whether or not to adopt the rules, which seems to require comparisons of what would happen if we adopted the rules and what would happen otherwise.

But assume that the Patterned View could avoid these problems. Even so, we should still consider the challenge in its strongest form. This justifies simply stipulating that the patronage system will last until the end of time – it *could* be replaced with a merit-based system, but it *won't*. Common-sense and consequentialist intuitions would still suggest that the patronage system *should* be replaced with a merit-based system. The mere fact that a convention is persistent does not justify it, so long as better alternatives are available.

4 Conclusion

Hodgson and Feldman presented forceful objections to early versions of actual rule consequentialism – particularly that such theories are too conservative, in failing to require individuals to actively improve their society's conventions, and objectionably conventionalist, in sanctioning immoral conventions. In this paper, I have argued that more recent versions of actual rule consequentialism continue to face these same problems.

Is there a more general lesson to be learned? I think so. Compared to ideal rule consequentialism, all versions of actual rule consequentialism restrict the scope of rules that can be evaluated as optimal and thus authoritative. The restriction might be to rules that are actually adopted, as in Miller's case, or to rules that relevantly correlate with actual patterns of behavior, as in Perl's case. By contrast, ideal rule consequentialism can evaluate all the rules sanctioned by actual rule consequentialism as well as others. These other rules will often be preferable, both according to common-sense moral intuitions

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

and from a generic consequentialist point of view. This, I suggest, is what makes actual rule consequentialism unconvincing.

Acknowledgements. Previous versions of this article were presented at LMU Munich, the University of Helsinki, and Linköping University. I would like to thank the audiences for their feedback.

Financial support. Financial support from Riksbankens Jubileumsfond (P22-0662) is gratefully acknowledged.

References

Bailey, James Wood. *Utilitarianism, Institutions, and Justice* (Oxford University Press, 1997).

Bykvist, Krister. *Utilitarianism: A Guide for the Perplexed* (Continuum, 2009).

Feldman, Fred. *Introductory Ethics* (Prentice-Hall, 1978).

Friedman, Ray, Ying-Yi Hong, Tony Simons, Shu-Cheng (Steve) Chi, Se-Hyung (David) Oh, and Mark Lachowicz. 'The Impact of Culture on Reactions to Promise Breaches: Differences between East and West in Behavioral Integrity Perceptions', *Group & Organization Management*, **43** (2) (2018), 273–315. https://doi.org/10.1177/1059601116678101.

Gauthier, David P. *Morals by Agreement* (Oxford University Press, 1986).

Hardin, Russell. *Morality within the Limits of Reason* (University of Chicago Press, (1988).

Harrod, Roy Forbes. 'Utilitarianism Revised', *Mind*, **45** (178) (1936), 137–56. https://doi.org/10.1093/mind/xlv.178.137.

Harsanyi, John C. 'Morality and the Theory of Rational Behavior', *Social Research: An International Quarterly*, **44** (4) (1977a), 623–56.

Harsanyi, John C. 'Rule Utilitarianism and Decision Theory', *Erkenntnis*, **11** (1) (1977b), 25–53. https://doi.org/10.1007/bf00169843.

Harsanyi, John C. 'Expectation Effects, Individual Utilities, and Rational Desires', in *Rationality, Rules and Utility: New Essays on the Moral Philosophy of Richard B. Brandt*, eds. by Brad Hooker (Westview Press, 1993), 115–126.

Hodgson, David. *Consequences of Utilitarianism: A Study in Normative Ethics and Legal Theory* (Clarendon Press, 1967).

Hooker, Brad. 'Ross-Style Pluralism Versus Rule-Consequentialism', *Mind*, **105** (420) (1996), 531–52. https://doi.org/10.1093/mind/105.420.531.

Hooker, Brad. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Clarendon Press, (2000).

Hooker, Brad. 'Rule Consequentialism', in *The Stanford Encyclopedia of Philosophy, Spring 2023*, eds. by Edward N. Zalta and Uri Nodelman (Metaphysics Research Lab, Stanford University, 2023). https://plato.stanford.edu/archives/spr2023/entries/consequentialism-rule/.

Johnson, Conrad D. *Moral Legislation: A Legal-Political Model for Indirect Consequentialist Reasoning* (Cambridge University Press, 1991).

Knight, Carl. 'Reflective Equilibrium', in *The Stanford Encyclopedia of Philosophy, Spring 2025*, eds. by Edward N. Zalta and Uri Nodelman (Metaphysics Research Lab, Stanford University, 2025). https://plato.stanford.edu/archives/spr2025/entries/reflective-equilibrium/.

Ma, Guoxin, John Blenkinsopp, and Steve Armstrong. 'Same Words, Different Worlds: Exploring Differences in Researcher and Participant Understandings of Promise and Obligation in the Psychological Contract', *Journal of Management & Organization*, **26** (1) (2020), 95–109. https://doi.org/10.1017/jmo.2018.76.

Mill, John Stuart. *Utilitarianism*. ed. by Roger Crisp (Oxford University Press UK, [1861]1998).

Miller, Richard B. 'Actual Rule Utilitarianism', *Journal of Philosophy*, **106** (1) (2009), 5–28. https://doi.org/10.5840/jphil200910611.

Mulgan, Tim. *Understanding Utilitarianism* (Routledge, 2007).

Perl, Caleb. 'Solving the Ideal Worlds Problem', *Ethics*, **132** (1) (2021), 89–126. https://doi.org/10.1086/715289.

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press

Perl, Caleb. 'Some Question-Begging Objections to Rule Consequentialism', *Australasian Journal of Philosophy*, **101** (4) (2023), 904–19. https://doi.org/10.1080/00048402.2022.2056213.

Podgorski, Abelard. 'Wouldn't It Be Nice? Moral Rules and Distant Worlds', *Noûs*, **52** (2) (2018), 279–94. https://doi.org/10.1111/nous.12189.

Rabinowicz, Wlodek. 'Does Practical Deliberation Crowd Out Self-Prediction?' *Erkenntnis*, **57** (1) (2002), 91–122.

Rawls, John. 'Two Concepts of Rules', *Philosophical Review*, **64** (1) (1955), 3–32. https://doi.org/10.2307/2182230.

Rumbold, Benedict. 'Does the Patterned View Avoid the Ideal Worlds Objection?' *Utilitas*, **36** (2) (2024), 130–47. https://doi.org/10.1017/s0953820824000037.

Shaw, William H. *Contemporary Ethics: Taking Account of Utilitarianism* (Wiley-Blackwell, 1999).

Slote, Michael and Philip Pettit. 'Satisficing Consequentialism', *Aristotelian Society Supplementary Volume*, **58** (1) (1984), 139–76. https://doi.org/10.1093/aristoteliansupp/58.1.139.

Cite this article: Andrić, Vuko. 2026. Enduring Objections to Actual Rule Consequentialism. *Utilitas* **38**, 289–302. https://doi.org/10.1017/S0953820826100429

https://doi.org/10.1017/S0953820826100429 Published online by Cambridge University Press