https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

RULE CONSEQUENTIALISM MAKES SENSE AFTER ALL

By Tyler Cowen

I. Introduction

Should we analyze policies by asking how they fit into more-general rules? For instance, maybe a single act of corruption has no harmful effects but corruption in general is harmful and many corrupt acts will destroy a polity. Would we be justified in condemning the single act of corruption?

More generally, does a "rules-based perspective" on policy evaluation possess independent force? Or does the doctrine of "rule consequentialism" collapse into "act consequentialism"? Why not make an exception to the rule when the exception is harmless or possibly even beneficial? Defenders of the rules-based perspective are wary of such questions. They know that if we are not willing to adhere to a rule for the sake of the rule itself, then talk of the rule is no better than a myth, albeit a socially useful myth.

Much is at stake. If rules-based perspectives hold up, it is easier to justify strict moralities and strict policy prescriptions and to make a much tighter, more consistent case for our preferred vision of a liberal order. We need only argue that the rule makes sense as a whole, not that it gives the best result in each and every case. Adopting a rules-based perspective also would make us less pragmatic and more willing to generalize. If the generality were true (e.g., "Don't gamble!"), that would be enough to establish its force, even if it didn't offer the best advice for every particular case. The rules-based approach mimics some of the properties of a rights theory, while allowing the rules to be determined by an empirical examination of what is beneficial to human society. It raises the possibility of true foundations for our vision of the good society.

A single act of weakness may not damage credibility much, but many weak acts will destroy credibility. A single investment in research may not boost science very much, but a broader collection of such investments will make a big difference. A single bad tax may not stifle economic growth, but a large number of bad taxes will. Numerous legal "exceptions" may make sense on an individual basis, but only if we fail to consider the more general consequences of those exceptions; for instance, many arguments for the rule of law depend on expectation and incentive effects. In general, when expectations and incentives matter, a rules-based approach and a case-by-case or act-based approach will offer different normative recommendations.

doi:10.1017/S0265052510000270

© 2011 Social Philosophy & Policy Foundation. Printed in the USA.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

Or consider the now-dormant Waxman-Markey cap-and-trade legislation (and related ideas), which would limit carbon emissions in the United States. Under one view, this is a proposal with "some cost and no real benefit"; the policy outlined in the bill will not lower global temperatures very much, in part because so many other countries emit carbon as well. Under another view, the bill is "the first step along a much-needed program to solve the dire and critical problem of climate change." Which view we accept again depends on whether we view the legislation in stand-alone or in bundled terms. The case for the bill is stronger if we adopt the bundled perspective (and, of course, if we see limiting climate change as a priority). In stand-alone terms, it is hard to justify the bill, as it costs something and achieves virtually no good in terms of its net, stand-alone impact on global climate.

The issue of time consistency also creates differences between act-based and rules-based perspectives. Time consistency logic covers a wide range of practical problems, including taxation, monetary policy, trade policy, patents, copyright, research and development, and foreign policy. To give one example, *ex ante* it is optimal to promise not to break the patent rights for a valuable new drug. *Ex post* it is optimal for the policymaker to confiscate the rights to the new drug once it is made. Pharmaceutical companies, of course, understand this logic, and they are somewhat reluctant to invest in new drugs in the first place. We would be better off if a no-confiscation rule could be put in place, but once the rule is in place we are better off if we can break it. That is a classic problem of time consistency.[1]

Some of the relevant examples of bundling are about bundling over time (time consistency), whereas others are about bundling across the actions of many different individuals. But in both cases we have the essential feature of doing normative evaluation while putting "up for grabs" more than the single choice of a lone individual.

As this essay proceeds, I will use the term "bundling" for when we treat choices as a collective package, and I will refer to the "small unit perspective" when we evaluate the choices in terms of their smaller components, such as the participation of a single additional person in the firing squad. Those are simply more general terms for the more commonly described notions of rule-based and act-based perspectives. If you wish, you also can substitute the more philosophically familiar terms "rule utilitarianism" and "act utilitarianism." The points here apply to a broader context than utilitarianism alone, but the logic of the issues very much resembles the classic rule utilitarian versus act utilitarian debates, even though bundling is the most fundamental concept to this inquiry.

[1] See Finn Kydland and Edward Prescott, "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85, no. 3 (1977): 473–92, among many others. Amihai Glazer and Lawrence S. Rothenberg outline policy issues involving time consistency issues in *Why Government Succeeds and Why It Fails* (Cambridge, CA: Harvard University Press, 2005).

One of Derek Parfit's "Mistakes in Moral Arithmetic," as presented in his book *Reasons and Persons* (1984), outlines a classic divide between an act-based and a rules-based perspective.[2] If a firing squad of six shooters kills an innocent person, all of them firing accurately at his heart, can we say that any one of the shooters is a murderer? After all, the "marginal product" of any single shooter was zero. Should we punish or invest resources to prevent any one of the potential shooters? Does it matter whose bullet arrived first? Should we refuse to prosecute group murders of this kind? How would we feel about a marksman who participates in a group hit on an innocent man, to earn a few dollars, figuring that the fellow is going to die anyway? This marksman might even donate some of those dollars to charities, or to nonprofits that fight to repeal capital punishment. An act-based perspective may suggest that participation in the shooting is fine, but a rules-based perspective creates room for strong objections.[3]

When our actions have nonlinear consequences, much is at stake when we choose whether to bundle and what to bundle. If we bundle the various marksmen's shots and evaluate the overall package of actions, participation will be understood as heinous and destructive, even if a particular marksman is not decisive with regard to the death of the victim and even if the marksman donates his wages to charity. The action, considered in a broader, bundled perspective, causes the death of an innocent human being. Yet if we stick with the small unit perspective, the marksman is improving social welfare. He is killing no one at the specified margin of choice, and both he and the charity may be better off.[4]

F. A. Hayek, James Buchanan, and Richard Epstein have, to varying degrees and in varying manners, linked the case for a free society to the validity of a rules-oriented perspective.[5] The idea is that we should look for those rules which will best delineate the scope of government, rather than evaluating all government policies on a case-by-case basis. In general, a rules-based perspective makes it easier to advocate or defend relatively extreme positions. Exceptions, by their very nature, are mod-

[2] Derek Parfit, *Reasons and Persons,* rev. ed. (Oxford: Clarendon Press, 1987).

[3] Ibid., 67, provides a string of related examples and conundrums.

[4] Some of these dilemmas resemble the Sorites problem, which typically involves nonlinear effects. A classic example of the Sorites problem is to ask how many stones constitute a pile. The contribution of any single stone to the "pileness" of the pile is zero or very small, yet the accretion of successive stones brings a pile into being. The analogy is not perfect, because our definition of "pile" is fuzzy, a complication which does not arise in the firing squad case (the death of the victim is unambiguous). Larry S. Temkin considers how the Sorites problem differs from intransitivity and vagueness as issues in moral philosophy in "A Continuum Argument for Intransitivity," *Philosophy and Public Affairs* 25, no. 3 (1996): 175–210.

[5] An extensive literature covers the practical arguments in favor of rules. See Friedrich A. Hayek, *The Constitution of Liberty* (Chicago: University of Chicago Press, 1960); Geoffrey Brennan and James M. Buchanan, *The Reason of Rules* (Indianapolis, IN: Liberty Fund, 2000); and Richard Epstein, *Simple Rules for a Complex World* (Cambridge, MA: Harvard University Press, 1997).

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

erating, relative to the alternative stance of following the rule all of the time. For instance, a rule that says "Don't bail out failed automobile companies" might be a good idea to prevent indiscriminate bailouts, but it also might be better to break that rule in particular circumstances. Under these assumptions, we will advocate more bailouts if we abandon the rules-based approach for an act-based perspective.

A rules-based perspective may seem attractive, but it has been on the intellectual defensive. To many observers, rule consequentialism is not persuasive, precisely because the underlying concern with consequences leads us to make many exceptions to the rule. If consequences are what matter, then adherence to the rule—for the rule's sake—is hard to justify. It seems we should break the rule when that would bring better consequences.

The strongest arguments against bundling recur in the debate between act and rule utilitarianism. Although I do not wish to defend the exclusive focus on utility found in many of those debates, I will at times borrow the terminology of that literature for expository purposes.

The act utilitarian typically believes that the rule utilitarian does not offer a coherent point of view. Act utilitarians (and others) apply the following *reduction* to rule utilitarianism: "What we really care about is utility (or some other notion of good consequences), not rules per se. So we should act to maximize utility. Sometimes adhering to rules maximizes utility, and then we adhere to rules, but sometimes breaking the rules maximizes utility. So rule utilitarianism, upon scrutiny, ought to collapse into act utilitarianism." Many philosophers accept this "reduction argument," and thus they do not consider rule utilitarianism (or rule consequentialism) to be an independent alternative to an act-based orientation. In debate, it is hard to beat the simple and immediately forceful reduction argument. After all, why *shouldn't* we make an exception and do something which will improve human welfare? The act-based orientation, of course, corresponds to what I am calling the small unit perspective.[6]

These issues have been debated since at least the eighteenth century, but I will propose a new path forward. I will address the oft-cited but underanalyzed concept of *feasibility*. I will start with the idea of trying to define the "feasible set"—or, alternatively, to define "constraints"—in a

[6] The literature here is enormous. See, for instance, Richard Brandt, "Toward a Credible Theory of Utilitarianism," in Hector-Neri Castaneda and George Nakhnikian, eds., *Morality and the Language of Conduct* (Detroit, MI: Wayne State University Press, 1963): 107–43; David Lyons, *Forms and Limits of Utilitarianism* (Oxford: Clarendon Press, 1965); Donald H. Regan, *Utilitarianism and Cooperation* (Oxford: Oxford University Press, 1980); Michael Slote, *From Morality to Virtue* (Oxford: Oxford University Press, 1992); Brad Hooker, *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Oxford: Clarendon Press, 2000): 93–111; John Leslie Mackie, *Persons and Values: Selected Papers, Vol. II* (Oxford: Clarendon Press, 1985); Geoffrey Scarre, *Utilitarianism* (New York: Routledge, 1996); and Fred Feldman, *Utilitarianism, Hedonism, and Desert* (Cambridge: Cambridge University Press, 1997), among many others. This problem has been present in rule utilitarianism since William Paley in the eighteenth century; see J. B. Schneewind, *Sidgwick's Ethics and Victorian Moral Philosophy* (Oxford: Clarendon Press, 1977), 125–27.

maximization problem in a philosophically coherent way. This will turn out to be a difficult problem, and it also will turn out to have some connections with the differences between bundled and small unit perspectives. This way of posing the question will give us access to some new intuitions. It turns out that our best current understanding of the concept of a "constraint" does not boil down to a simple matter of fact. Rather, it involves the murky waters of modal logic—namely, the philosophical literature on what it means to postulate that things might have been different from what they are or were. As it will turn out, the fuzziness of modal considerations will lead us to a partial license for rules-based thinking, yet without a rules-based perspective collapsing into an act-based perspective.

II. What Is Feasible and What Is Utopian?

If we think of utilitarian or consequentialist doctrine as involving problems of constrained maximization, we might look to economics for clarification on what it means to postulate a constraint. At the textbook level, economists use the idea of a budget constraint to delineate the utopian from the feasible. In this view, "moving along the budget constraint" (reshuffling resources) is feasible, whereas "wanting the budget constraint to shift out" (i.e., wanting more resources for no cost) is considered to be utopian, akin to wishing for the proverbial free lunch.

Here is a simple budget constraint:

$$\text{Maximize U, subject to } p1x1 + p2x2 = Y$$

When we graph this constraint, we get the familiar picture in figure 1.

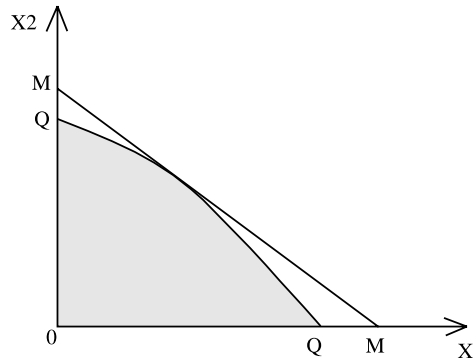


Figure 1. Production possibilities frontier QQ cannot exceed budget constraint MM. The shaded area represents feasible combinations of x1 and x2, given total wealth and resource transformation capabilities.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

The diagram in figure 1 appears to indicate that there is a clear difference between "moving outside the constraint" and "moving along the constraint," as illustrated by the lines and descriptions in the diagram. That is typically how economists define what is feasible and what is not.

This distinction, however useful it may be to positive or predictive economics, begs the question as to what is truly feasible at a more fundamental level. A society cannot move from one point along a budget constraint to another point without a cost being incurred somehow, somewhere. The resources expressed and measured by the budget constraint are all owned and controlled by various individuals and, in the absence of interference, agents will allocate these resources one way rather than another. To ask for one allocation rather than another is to stipulate that some existing constraints and incentives be changed. In other words, when moving along the budget constraint, we also are asking for more resources, or for different resources, albeit in disguised fashion. At the very least, we are asking for something to be different, relative to an observed status quo. We are not involved in any straightforward process of choosing among options which "lie before us," like various pairs of socks might be laid out on a bedspread each morning when one is dressing for work. Bringing about other options along the drawn budget constraint requires different facts associated with a different possible world.

Sometimes economists think about this dilemma in terms of transaction costs. When drawing a social budget constraint, we are missing the transaction costs dimension from our measurement of costs. In considering a move along the budget constraint, we are wishing that outcomes could take a different particular form or, in other words, that transactions costs in some manner would be different. We are imagining the involved agents cutting a different set of deals, rather than the status quo. That's another way of wishing for a change, and, on reflection, it is not obviously different from wishing that manna would fall from the sky and wishing to live "beyond" the budget constraint. The specification "Transactions costs change, so we can move along the budget constraint" slides all too easily into "Lower costs would bring a better outcome and push us outside the previous budget constraint."

In any case, the real question is what the social budget constraint looks like in the first place. Once we view it in these terms, we can no longer invoke the budget constraint as an *a priori* circumscription of what is feasible and what is not. Standard economic methods simply assume that we have defined the feasible, rather than providing such a definition.

Milton Friedman, in an essay for the free-market Cato Institute, laid out the dilemma very clearly, albeit unintentionally.[7] He titled his essay "The

[7] Milton Friedman, "The Real Free Lunch: Markets and Private Property," in David Boaz, ed., *Toward Liberty: The Idea That Is Changing the World* (Washington DC: Cato Institute, 2002), 55–62.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

Real Free Lunch: Markets and Private Property," and he argued that market economies capture gains from trade and can make virtually everyone better off, if only we would rely on markets more. Whether you agree with Friedman's positive claim about markets is not the point for this discussion. Rather, I am interested in how he imagined the feasible set. Had Friedman not insisted elsewhere that "There Is No Such Thing as a Free Lunch"? Is not asking for "more markets" another kind of utopianism, no more relevant than the plea for a free lunch?

"What if" we all had more information? What if, to paraphrase James Madison, men were angels? Would not the world be better off? In each case the answer is yes, or could be yes if we worded the query to cover the relevant counterexamples. Yet the real question is what conclusions we can draw from such simple comparisons. Economists are renowned for refusing to claim that the absence of a free lunch represents inefficiency per se.[8] Simply listing some states of affairs which would be better does not establish a meaningful suggestion for political policy or reform. It also would be better if we didn't have earthquakes.

In denying the normative relevance of the free lunch, economists are staking out (part of) a position on feasibility, if only implicitly. Let us see how that position looks if we present it in terms of a simple ordering. Economists, for instance, might rank some alternatives, in terms of their degree of feasibility, in the following manner:

Most to least utopian:

(1) Oceans of lemonade
(2) Social state X plus a free lunch
(3) Social state X (with different policies than the status quo)
(4) Status quo

Economists, in denying the relevance of the free lunch, are unwilling to criticize vision (3) for falling short of vision (2); in other words, vision (2) is excessively utopian, or (alternatively) it is not sufficiently feasible.

At the same time, many economists do not see it as excessively utopian to ask for (3) in lieu of (4). In lieu of making pleas for free lunches, many contemporary economists advocate what they call "comparative institutional analysis." The economist compares one set of (feasible?) institutions to another set of (feasible?) institutions and asks which is better. We have a relevant imperfection only when an institutional improvement would bring more benefit than cost.

In other words, many economists consider it normatively interesting to ask:

[8] Harold Demsetz, "Information and Efficiency: Another Viewpoint," *Journal of Law and Economics* 12, no. 1 (1969): 1–22.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

(1) What if we had institution X instead of institution Y?

Those same economists do not consider it normatively interesting to ask:

(2) What if we had a free lunch?

The distinction between these two questions is not easy to defend on the grounds of pure logic. It is difficult to dismiss the free lunch question as utopian while maintaining the normative relevance of the comparative institutional question. If institution X is better than institution Y, then arguably having X is one form of a free lunch. Preferring institution X over institution Y could be, to make a simple comparison, as utopian as wishing the world didn't have major earthquakes or that human beings were more benevolent.[9]

Economists therefore face a "war on two fronts," to borrow Parfit's fruitful concept.[10] It is not easy to reconcile two strong intuitions. The first is that inefficiencies exist and can be pinpointed meaningfully. The second is that there is no free lunch, and it is normatively meaningless to wish for one. Either intuition, taken alone, sounds plausible. But it is hard to put the two intuitions together into a single account of which changes are feasible and which are not.[11]

Moving outside economics to the context of political thought, we can find similar issues in different contexts. A world without scarcity may sound good as an abstract ideal, but we reject the idea as excessively utopian. Fair enough, but now consider some growth-improving policy, such as a revamping of the U.S. tax system. Given all the special interest groups at work, is wishing for a truly good reform much more realistic than Charles Fourier's belief that socialism would bring oceans of lem-

[9] This ranked example is drawn from Tyler Cowen, "The Epistemic Problem Does Not Refute Consequentialism," *Utilitas* 18, no. 4 (2006): 383–99. On the concept of utopianism in the economics literature, see Clarence Philbrook, "'Realism' in Policy Espousal," *American Economic Review* 43, no. 5 (1953): 846–59, reprinted in Daniel B. Klein, ed., *What Do Economists Contribute?* (Washington, DC: Cato Institute, 1999), 69–86; Carl J. Dahlman, "The Problem of Externality," *Journal of Law and Economics* 22, no. 1 (1979): 141–62; Pamela J. Brown, *Deadweight Loss: A Nonexistence Theorem* (unpublished manuscript, California State University, Northridge, 1988); and Daniel B. Klein, ed., *What Do Economists Contribute?* (Washington, DC: Cato Institute, 1999).

[10] Parfit, *Reasons and Persons,* Part I, chap. 3.

[11] Alastair Norcross argues that we need to consider the best available action relative to alternatives, and discusses the ambiguities in defining exactly what those alternatives are, in "Good and Bad Actions," *The Philosophical Review* 106, no. 1 (1997): 1–33. On the relevance of related ideas for the free will controversies, see Daniel C. Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press, 1984). For discussion of the meanings of "if" and "can" in ordinary language philosophy, see J. L. Austin, *Philosophical Papers* (Oxford: Clarendon Press, 1961); and D. F. Pears, "Ifs and Cans," in *Essays on J. L. Austin* (Oxford: Clarendon Press, 1973). Simon Blackburn considers general issues involving morals and modal logic in "Morals and Modals," in Blackburn, *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993), 52–74.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

onade and ship-pulling dolphins?[12] It might be more realistic as a matter of degree, but it is hard to see where a clear line—delineating the feasible from the infeasible—might come from.

The general issue is this. We wish to allow at least *some* scope for normative, big-picture thinking. It is not useful to argue that the world is already as good as it could be. At the same time, we do not wish to think so big that we lose the moorings of common sense and veer off into the excessively utopian. It doesn't make sense to just draw up an ideal wish list and call that an approach to reform.

Many reform advocates wish, at least implicitly, to have it both ways. They require that some degree of utopianism is acceptable. Proposals for change, to some extent, require that we advocate good outcomes for their own sake, without necessarily predicting their adoption. Without this willingness to be at least somewhat "utopian," we cannot elevate a good reform proposal above the status quo. If we advocated only what was already in place, expected, or forthcoming, there would be little point to making policy prescriptions. Nonetheless, for whatever imagined improvement we can cite, there will exist better, yet more utopian, competing proposals. We reject (many of these) proposals by arguing that they are excessively utopian. Yet we again return to the dilemma: Why do we classify some reform proposals as feasible and others as infeasible?[13]

The standard "conservative" critiques of utopianism, as we find in Karl Popper, F. A. Hayek, and evolutionary biology, do not much help in this context. These powerful arguments do tell us that human nature is somewhat fixed or even extremely fixed; they suggest that certain proposals, especially of a totalitarian nature, would be disasters rather than utopias.

[12] In fairness to Fourier, he was also an early prophet of the steam locomotive, a view for which he was ridiculed in his time; see Jonathan Beecher, *Charles Fourier: The Visionary and His World* (Berkeley: University of California Press, 1986), 59. On the lemonade idea, see ibid., 125.

[13] The literature on utopias raises related questions, although not usually in a philosophic or rational choice framework. Aurel Kolnai writes: "How exactly can we distinguish between the proper pursuit of the good and its perfectionist aberration?" in Kolnai, *The Utopian Mind and Other Papers: A Critical Study in Moral and Political Philosophy,* ed. Francis Dunlop (London: Athlone, 1995), 17. Frank E. Manuel and Fritzie P. Manuel note that "one man's trivial revision is another man's upheaval" in *Utopian Thought in the Western World* (Cambridge, MA: Belknap Press of Harvard University Press, 1979), 8. Karl Mannheim refers to the "difficulty in defining precisely what, at a given period, is to be regarded as ideology, and what as utopia" in *Ideology and Utopia: An Introduction to the Sociology of Knowledge* (New York: Harcourt, 1936), 203. Since at least Friedrich Engels, this topic has been a staple of socialist debate as well. See Friedrich Engels, *Socialism: Utopian and Scientific* (Peking: Foreign Languages Press, 1975). For a survey of some definitions of utopia, see Ruth Levitas, *The Concept of Utopia* (Syracuse, NY: Syracuse University Press, 1990), 3. See also J. C. Davis, "The History of Utopia: The Chronology of Nowhere," in Peter Alexander and Roger Gill, eds., *Utopias* (London: Duckworth, 1984): 1–18; and Lyman Tower Sargent, "Utopian Themes: Themes and Variations," in Roland Schaer, Gregory Claeys, and Lyman Tower Sargent, eds., *Utopia: The Search for the Ideal Society in the Western World* (New York: The New Public Library/Oxford University Press, 2000), 8–15. Barbara Goodwin and Keith Taylor consider the role that concepts of utopia have played in political debate in *The Politics of Utopia: A Study in Theory and Practice* (New York: St. Martin's Press, 1982).

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

But these perspectives do not address the utopian dilemma per se, much less resolve it. The key question is not about central planning. Rather, it is about whether, when we consider changes (including a move away from central planning), we should evaluate those changes bit by bit or together as part of a broader bundle. In other words, invoking Popper, Hayek, or conservative thinkers does not settle what is a metaphysical rather than a practical planning question. Indeed, as mentioned above, Hayek himself advocated a high degree of rules-based thinking.

We also cannot look to formal modal logic to resolve these questions about feasibility. Formal modal logic is a well-developed philosophic literature which looks at what it means to analyze or speak of "possible worlds." Nonetheless, without intending any criticism of the broader genre, modal logic is not well-suited to the task at hand. First, the major approaches to modal logic deal with a very broadly conceived notion of what is feasible or possible. For instance, it is frequently accepted that "talking donkeys," however strange the concept may be in common-sense terms, belong to the set of possible worlds. Maybe such donkeys are not impossible in the sense of violating the known laws of physics, but arguably they are still "too impossible" for the more practical and applied purposes of ethics and politics, where far less absurd ideas are dismissed out of hand. Modal logic operates within a broader notion of the feasible than would resolve typical policy debates over the feasible set. I am looking for a concrete and practically useful method of judging feasibility, rather than a purely abstract standard of logical classification based on hard constraints taken from the science of physics.

Second, modal logic itself presents numerous unsolved dilemmas at a quite fundamental level, not the least of which is what the concept of "possible worlds" means in the first place. Some philosophers believe that claims about modal logic make sense only if they refer to real worlds actually out there (modal realism), while other philosophers are suspicious of the concept of modality altogether. I think of the literature on modal logic as reflecting and anticipating the more practical dilemmas discussed above, not as solving them. Ideally, we would like a solution, or at least a means of moving forward, which is independent of any very particular approach to controversial questions of metaphysics and modal logic. In any case, we still must plow forward using stop-gap approaches in the meantime.[14]

[14] On various modal debates, see Michael J. Loux, *The Possible and the Actual: Readings in the Metaphysics of Modality* (Ithaca, NY: Cornell University Press, 1979); Graeme Forbes, *The Metaphysics of Modality* (Oxford: Clarendon Press, 1985); David Lewis, *On the Plurality of Worlds* (New York: Basil Blackwell, 1986); D. M. Armstrong, *A Combinatorial Theory of Possibility* (Cambridge: Cambridge University Press, 1989); William G. Lycan, *Modality and Meaning* (Dordrecht: Kluwer Publishers, 1994); Christopher Hitchcock, "Farewell to Binary Causation," *Canadian Journal of Philosophy* 26, no. 2 (1996): 267–82; Alexander Robert Pruss, "Possible Worlds: What They Are Good For and What They Are Not" (unpublished doctoral dissertation, University of Pittsburgh, 2001); Theodore Sider, "The Ersatz Pluriverse," *The*

III. The Panglossian View as a Stopping Point

Some economists talk as if they believe the extreme claim that all observed outcomes are efficient. Nobel laureate George Stigler and other members of the Chicago School have been associated with this view, although it does not seem that Stigler ever embraced it in print. It may be more of a lunch-room exercise than a view that anyone holds consistently, but I am interested in it nonetheless. Sometimes this claim is presented in a modified form, such as "Everything we observe in markets is efficient," "Democracies are efficient," or "Everything is locally efficient, albeit not always globally efficient." I will consider the most straightforward claim that everything observed is efficient, all of the time. It might seem that I am picking on this attitude by examining its weakest variant, but my purpose is the opposite. I would like to argue that the (ostensibly) weakest version of this view is, in terms of logic, stronger and more plausible than many of the more popular intermediate positions.[15]

Although virtually everyone rejects this Panglossian view, it has a logical consistency which many of its competitors do not. Any beneficial improvement which we do not already have is utopian and thus should be dismissed as a free lunch. In essence, the budget constraint is now a single point, namely the status quo.

One who holds the Panglossian view has a rejoinder to claims of inefficiency: "The current state of affairs *would* be inefficient, if the relevant parties could bargain or trade to bring about a better outcome. But apparently they cannot. Correcting the so-called problem is too costly. The existence of the problem is efficient, once we take *all* constraints and *all* costs, including the costs of bargaining and the costs of change, into account. To claim otherwise is simply to wish that things would be better, a kind of utopian dreaming. We are economists. We do not ask for free lunches."

Of course, the Panglossian view need not be thought of as especially optimistic. We are in the "worst of all possible worlds" as well as in the "best of all possible worlds."

At this point, many individuals will mock the Panglossian view as a tautology or a moral absurdity, more deserving of scorn than an intellectual response. Genocide happens, babies starve, the mail is not always delivered on time, and so on. Surely something is wrong in the world,

Journal of Philosophy 99, no. 6 (2002): 279–315; Tamar Szabo Gendler and John Hawthorne, eds., *Conceivability and Possibility* (Oxford: Clarendon Press, 2002); and John Divers, *Possible Worlds* (London: Routledge, 2002).

[15] Parmenides argued that the world could not be any different. The literature on theodicy considers whether God made the "best possible world" and what it means to say that other worlds are possible. See Robert M. Adams, *The Virtue of Faith and Other Essays in Philosophical Ideology* (New York: Oxford University Press, 1987); Robert M. Adams, *Leibniz: Determinist, Theist, Idealist* (New York: Oxford University Press, 1994); and Alvin Plantinga, *The Nature of Necessity* (Oxford: Clarendon Press, 1989).

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

and it outrages many people to hear that everything is efficient, all things considered. The Panglossian view seems to abuse discourse and our common-sense understanding that real-world improvements are possible.

By the way, the Panglossian view is only one of the available extremes. John Stuart Mill, Marquis de Condorcet, and Herbert Spencer all defended the perfectibility of mankind as a central political vision. Today, tech-savvy advocates of "The Singularity" predict that we will be genetically reengineered for the better or we will become computer uploads. These attitudes, like the Panglossian view, offer a logical consistency of sorts. They tell us to hold out for an extreme utopian vision in which virtually everything is up for grabs and very broad improvements will prove feasible.[16]

Most of us embrace some intermediate view between Panglossianism and extreme utopianism. An intermediate view "feels" right, seems reasonable, and fits with the practice of most of the smart and honest people we know. Yet the intermediate positions, appealing though they may sound, do not have as much logic on their side as we might like to think. Once you push on them, it is hard to defend why we might allow one degree of utopianism yet reject another. The intermediate positions run a high risk of being arbitrary rather than well-grounded and consistent.

The "marginalist" tradition, which in economics stresses looking at incremental changes, might lead many economists to favor the small unit perspective and to lean away from utopianism. Yet the positive analytical successes of economics, using the marginalist method, do not translate into a strong case for the small unit perspective as a means for distinguishing relevant moral comparisons from irrelevant moral comparisons. Once we see feasibility as a matter of degree, the relevant definition of "the margin" is not given immediately or unambiguously by the objective facts of the case.

Furthermore, the economics literature has a lot of well-regarded work which moves away from the small unit perspective. For instance, versions of the bundling problem surface repeatedly in game theory, but in that context the literature has not generally sided with the small unit perspective. One common construct is the "Shapley value," which is used to measure the importance of an individual to a broader coalition. A Shapley value may help measure a worker's contribution to joint production, or a voter's contribution to a political coalition. The details of the Shapley value need not concern us here, but the construct does not refer to an agent's marginal product in the traditional sense; rather, it refers to a metric of bundling. The Shapley value, while it hardly commands universal support, serves as the most popular "rational choice" solution to many bargaining problems. This suggests that economics and rational

[16] For a more general survey of utopian thinking, see Manuel and Manuel, *Utopian Thought in the Western World.*

choice approaches do not commit us to the small unit perspective when nonlinear effects are present.[17]

IV. Degrees of Feasibility

In lieu of the Panglossian and pure utopian points of view, we can think of feasibility as a spectrum rather than as an all-or-nothing category. Some specified world-states are more utopian than others, but as a matter of degree rather than of kind.

Most feasibility differences do appear to be differences of degree. For instance, consider the option of inducing more work from a labor force. Probably it is too utopian to expect everyone to work eighteen hours a day and enjoy it; human motivations would have to be "too different" for that to happen. But the judgment of nonfeasibility does not appear to kick in at any particular number of hours, as could be defined by an exact cut-off. Rather, the more extra work is specified, the lower the degree of feasibility that appears to hold. A similar logic holds with regard to many other feasibility judgments, including those in the areas of fiscal restraint, tax reform, changes in the speed of bureaucracy, or, for that matter, declines in the probability of genocide, referring to the world's more violent countries. In most cases, the larger the potential improvement, the more utopian it will appear to be, again without any clear cut-off point as to what is feasible and what is not.

At the very least, when we consider the practical epistemic issues, it is unlikely that we could identify or verify a cut-off point between the feasible and the infeasible, even if such a clear line existed in some rarefied realm of modal logic. Perhaps there is some ultimately true and correct metaphysical theory in which people can work seventeen minutes longer each, but not seventeen minutes and one millisecond longer. But again we are far from having access to such knowledge. For practical purposes, we are left with feasibility as a matter of degree.

We can see degrees of feasibility even in the Fourier vision of a utopian future. Forget about oceans, what if we were told that socialism would bring us a mid-sized lake full of lemonade? A small pond? How about a

[17] In formal terms, the Shapley solution looks at all possible differing "coalitions" (combinations of actions or abstentions from action, in the firing squad example) and measures the differing marginal values of an individual unit to the coalition. These marginal values are then averaged across all of the possible combinations of units. In the firing squad example, for instance, the Shapley value averages a single shooter's marginal impact across "all six of us shoot," "only the first five of the six shoot," "only the last five of the six shoot," "only these three of the six shoot," "only I shoot," and so on, across all the possible combinations. We will then find that the Shapley value for a single marksman is positive, but less than the value of an individual life. On the bargaining theory foundations for the Shapley value, see Alvin E. Roth, "Axiomatic Models of Bargaining," *Lecture Notes in Economics and Mathematical Systems,* No. 170 (New York: Springer Verlag, 1979); and Faruk Gul, "Bargaining Foundations of Shapley Value," *Econometrica* 57, no. 1 (1989): 81–95.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

glass full of lemonade? (Can you get that in North Korea? I wonder.) Again, it is difficult to find a distinct cut-off at which the specified world-states clearly and definitely cross from the realm of the feasible into the realm of the infeasible.

A spectrum of feasibility might specify a number of dimensions. A more feasible vision, compared to a less feasible vision, might be "more like" the world we know in terms of fact, more like the world we know in terms of adherence to the laws of science, or might stand at the end of a relatively simple path from "here" to "there." Most likely, the more feasible world-state possesses some combination of these qualities, relative to the less feasible world-state. For our purposes, the exact number or nature of the dimensions is not the relevant point; it suffices to note that some kind of continuum exists.[18]

In some cases physical laws might suggest very definite *limits* to feasibility, but still, along most margins, feasibility will be a matter of degree. Quantum theory, for instance, might suggest that particles can only become so small, at which point no further divisibility is possible. Perhaps we can measure this limit quite exactly in principle, if not always in practice. To take another example, in Einstein's theory no entities other than tachyons can travel faster than the speed of light, which again serves as a definite limit. But the feasibility problems of the social sciences are not so exact or so amenable to fixed measurement at these kinds of limiting points. We do not know how to break social science problems down into their quantum elements, and maybe it's not possible at all. We are left again with feasibility as a sliding scale, at least for the vast majority of our speculations, and our knowledge of physics offers few useful guidelines.[19]

V. In Defense of Rules-Based Thinking

Any plausible normative theory must, to some extent, treat some previously imagined constraints as freely floating parameters of choice. That is the take-away from the discussion of the Panglossian view. But we do not have clear guidance for where to stop in delineating the feasible set. In other words, there is no strong argument—apart from the Panglossian

[18] David Lewis has suggested some standards for ranking worlds in terms of their similarity, and along these lines we might regard the more similar worlds to our own as "more possible" or "less utopian." See David Lewis, "Counterfactual Dependence and Time's Arrow," *Noûs* 13, no. 4 (1979): 472. For a treatment of degrees of possibility, see Forbes, *The Metaphysics of Modality,* chap. 7.

[19] Note that we should not identify feasibility with the notions of probability or likelihood. Feasibility refers in some manner to the "closeness" of some other world to our own, whether or not we expect that world to occur. Blinking your eyes one more time in a day, each day, might be quite feasible in the common-sense use of that term, although we do not necessarily expect such an act to occur with a high probability. Conditional on the number of blinks changing, the chance that the change is exactly one blink might be quite small. This example suggests that feasibility and probability are distinct concepts and that a high degree of feasibility does not have to mean a high degree of probability.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

view—for keeping the degree of utopian thinking to a minimum, and such a view does not command much loyalty.

Rather than collapsing into act utilitarianism, rule utilitarianism specifies a differing series of constraints and thus a different maximization problem altogether. For instance, when it comes to bundling across time, we are picking some policy today, under the constraint that a similar policy (or rule) be applied in subsequent time periods. In contrast, act utilitarianism takes the action of one individual at one point in time as a free variable, and asks what that agent should do to maximize utility. Rule utilitarianism takes an entire series of individual actions to be free variables, but it considers them collectively rather than one-by-one. More is "up for grabs" than under act utilitarianism.

More formally, rule utilitarianism looks like this:

(1) Maximize U, subject to Action (t1) = Actions (t2, t3, t3. . .)

In other words, we have to apply a rule consistently over time, and across different persons, when it comes to chosen actions.

Act utilitarianism looks more like this:

(2) Maximize U, allowing for Actions (t1, t2, t3. . .) as different variables and selecting for t1 alone.

Under this understanding, the rules-based perspective is based on the notion of a difference in the constraints. Rule utilitarianism and act utilitarianism postulate different maximization problems, and thus they coexist in some fashion, rather than one subsuming the other or one collapsing into the other. Furthermore, since we do not have a definitive, fact-based way of presenting "the correct constraints," we do not have an in-principle argument for rejecting the rules-based perspective. There is still no argument that the rules-based perspective is better (more on that later), but at the very least it has no longer been knocked out of the running. The thing to be maximized (the maximand) is in both problems the same—some notion of good consequences, with utility as the immediate stand-in—but the relevant constraint is not uniquely determined in any manner which would allow us to adjudicate across the two problems.

For a simple analogy, consider two microeconomic problems. The first asks a business to maximize profit, subject to the constraint that inventories in one particular store never fall below a specified level. The second asks a business to maximize profit, subject to the constraint that inventories in each store in the chain never fall below a specified level. Again the maximand is the same—profit—but we have two differing constraints. It would be wrong to claim that one such problem is "the wrong normative theory," or that one such problem "collapses" into the other. Rule and act utilitarianism/consequentialism hold a similar dual and

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

possibly coexisting status, once we look at both the maximand *and the constraint*.

The reduction argument, as it is used against rules-based perspectives, requires a commitment to a very particular account of feasibility. It requires the implicit claim that thinking of rules or bundled acts as the relevant constraint is incorrect. Supposedly we should look at a smaller unit of choice—the single act—as the relevant constraint. But again, once we see feasibility as a sliding scale and a matter of degree, it is not obvious why this argument should have so much force. Treating "a bundle of choices" as a relevant free variable is no less defensible than treating "a single act" as a relevant free variable. If we can put a single act up for grabs in our normative decision-making, why not treat a larger rule or bundled package as a variable for choice?

The act-utilitarian critique of rule utilitarianism seems plausible because act utilitarianism appears to examine the smallest possible unit of choice, namely, the individual act. But the smallest possible unit of choice is *zero*—namely, no change—not a single act. The logic of the act-utilitarian reduction argument should, taken alone, lead to the Panglossian perspective, not to act utilitarianism. Act utilitarianism "survives," so to speak, only because we allow some utopianism in the door and allow at least one constraint to be loosened. If we can object to taking rules as a free-floating variable, someone who holds the Panglossian view can object to taking individual acts as a free-floating variable. Why stop at one point rather than the other? Once we elevate an "act" perspective over the Panglossian view, we have opened the door for yet additional flexibility in understanding the constraint, and the reduction argument is no longer decisive.

In other words, the act-utilitarian critique fails to win a "war on two fronts," again referring to Parfit's concept. The act-utilitarian perspective might appear strong when faced with either the Panglossian view or rule utilitarianism. But it has a harder time taking them both on at the same time and winning the debate. The logic of act utilitarianism—or, indeed, of normative analysis in the first place—allows for the possibility of moving even further away from the Panglossian view and considering rules-based changes or other bundled approaches to the constraints.

Having covered this ground, I would like to consider two issues which arise commonly in discussions of bundling or rules-based perspectives. One common criticism of the rules-based perspective is that the act-based perspective can embody quite complex instructions, sufficiently complex that rules are not needed. For instance, when coordination problems are present, we might need more complex instructions rather than very simple instructions. A simple instruction might be "Perform action X," whereas the more complex instruction might be "Perform action X, but only if enough other cooperators help make the collective effort worthwhile," or however the instructions might usefully be worded. The underlying claim

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

is that complex instructions are a more useful concept than rules, and that act utilitarianism, combined with sufficiently complex instructions, is robust. One can debate the semantic issue of whether this argument represents a victory for act or rule consequentialism or some other perspective. The important point is that the more complex instructions expand the scope of bundling beyond a single do/don't act. We are still left with narrower and broader notions of the constraints (no one conception of the constraints having a metaphysical priority over the others) and with some license for a broader notion of bundling.

A second common criticism is that a rules-based perspective cannot handle all possible counterexamples. Let us postulate a rule, namely, that everyone should pitch in to achieve some common end, such as moving a car out of the way, to open an important road for passage. It takes ten people to lift the car, and ten are present. We can all agree that all ten should follow the rule and lift the car. At the same time, if it is expected that three people will shirk their duties, and not lift, it makes no sense for the other seven to soil their clothes and strain their bodies, for no useful end. The rules-based perspective would appear to experience trouble with this example, since following a good rule with three noncooperators simply wastes an effort. But the correct conclusion is not that bundled perspectives fail. At best this example shows that there is no single, supreme rule. There still exists a series of imagined constraints—no behavior up for grabs, behavior from one of the agents up for grabs, seven behaviors up for grabs, three up for grabs, all ten up for grabs, and so on. Again, no one of those constraints has a privileged metaphysical position over the other. We might prefer to have "all ten" behaviors up for grabs as the most effective lever for manipulation. But those (hypothetical) practical benefits do not pin down or identify the appropriate level of feasibility or utopianness. To see why not, let's say we had a practical problem which required the tight coordination of billions of agents to be solved. That would not mean we had found "billions" to be an appropriate level of coordinative feasibility; at the same time, it also would not altogether rule out thinking in terms of broader bundles of choice or action.

VI. How Strong Can the Arguments for a Rules-Based Perspective Get?

So far, the argument has established only that the reduction argument, as I have called it, does not defeat a rules-based or bundled perspective. That is progress, but actually the arguments for bundling are stronger than that observation.

Most importantly, the arguments for bundling do not require bundling to be correct with certainty or with a probability of 1. As it stands, we have a set of coexisting perspectives about constraints and modality. One

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

possibility is that one specification of the constraint is shown to be correct and all the others will fall away; that's possible, but at this stage of the game it's not looking likely.

Another option is that all or many of the differing modal perspectives will remain on the table, and for practical choices we will have to add up or aggregate these differing views in some manner. The idea of coexisting moral codes is, in fact, a familiar one, and it appears in the debates on different versions of rule-utilitarian principles. For instance, there are multiple approaches to generating a rules-based code. Should we evaluate the code on the presupposition that it is adopted and internalized by all relevant individuals? Or should we evaluate the code assuming that we must bear significant costs of adoption and internalization? In other words, the code can be treated either as a free variable (to be selected or not), or as an investment in a new custom, including transition costs. Again, we see two maximization problems, each involving a different specification of constraints. Just as one problem does not "reduce" to the other, rules-based approaches do not reduce to the act-based standard. It also can be said that there is no obvious supreme rule, only different methods of thinking about bundling. In this regard, rules-based approaches have to make some very important concessions, and we cannot expect a rules-based approach to support a monolithic approach to moral thinking.[20]

That all said, when we have multiple, coexisting perspectives floating around in our moral theory, it remains true that in the world we must proceed to formulate some view, however provisional or understated. That suggests that we should look to some plausible aggregate or combination of those multiple perspectives, and if we take this approach, bundling has again survived the reduction argument. If we aggregate single-act and bundled perspectives, the overall blend will give some weight to bundling, by the very nature of aggregation. Ultimately that blend is more of a victory for the bundled perspective than for the single-act perspective.

For instance, take a policy that has no effect when applied in small doses but significantly boosts economic growth when it is applied repeatedly. Furthermore, assume that we believe bundling to be correct, applicable, or otherwise relevant or weight-worthy in some manner with, say, a probability of 0.2. In expected value terms, we still ought to pursue the recommendation of the bundled perspective, given how the calculation would proceed. This is a deliberately oversimplified example, but it reflects a more general point. If we combine rules-based and act-based perspectives, the very act of combining means that on net we are left with some form of bundling. More complicated examples will give weight to both

[20] On costs of adoption and internalization and rule utilitarianism, see Richard Brandt, "Toward a Credible Theory of Utilitarianism," in Castaneda and Nakhnikian, eds., *Morality and the Language of Conduct*, 107–43; and Hooker, *Ideal Code, Real World*, 78–79. On varieties of rule utilitarianism more generally, see Scarre, *Utilitarianism*, 122–32.

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

act-based and bundled perspectives, but in net terms some method of bundling, whatever the weights given to the bundled perspectives may be, will be influencing the final decisions.

Aggregation procedures are the friend of bundling, as a general method, not its enemy. In contrast, aggregation procedures move us away from the single-act perspective, which by its nature cannot be combined or weighted with anything else. So again, if we aggregate together bundled and single-act perspectives, the result generally will involve some bundling. That means the relevant question is not whether to bundle but rather which form or method of bundling is the best way to proceed. All of a sudden the debate has been settled in favor of bundling, in some form or another. In future work, currently in progress, I am pursuing the question of whether we can find a focal means of bundling, given the multiplicity of bundling options facing us.

When comparing more or less bundled alternatives, it is also worth asking the practical question: In which direction is our civilization most likely to err? Are we more likely to perish because our decision algorithms involved too much bundling or too little bundling?

In my view, it is unlikely that real-world decision-makers will bundle too much. The nature of politics is more likely to produce too little bundling than too much bundling. Policymakers often make decisions on a day-to-day, case-by-case basis, simply hoping to survive the next election cycle. Politicians typically look at the marginal products of particular actions—if indeed they are that sophisticated, rational, or public-minded in the first place—rather than treating a broader set of actions as a bundled group. That is one reason why we make as many political mistakes as we do.

To be sure, this kind of "noble lie" defense of a rules-based perspective does not get us very far on its own. For instance, plenty of act utilitarians will admit that we are better off for believing in rule utilitarianism, while still insisting that the latter cannot be philosophically justified. But the "noble lie" approach is too cynical a way of viewing the choice between rules-based and act-based perspectives. Calling it a noble lie is assuming that a rules-based perspective is a philosophical weak sister to begin with. I have tried to show it isn't. The notion of rules offers a legitimate and independent perspective which can compete with act consequentialism, or act utilitarianism, on at least equal terms and possibly superior terms.

If that is the case, the practical benefits of believing in a rules-based perspective do not have to be seen as a noble lie. They might just be part of a very important noble truth.

VII. Conclusion

The rules-based perspective deserves another look. The reduction argument, which is often used against a rules-based perspective, is not as

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press

strong as it looks. More generally, a lot in moral theory depends on how we think about defining the relevant constraints on our choices. To the extent that a rules-based perspective is defensible, it may be possible to revive a vision of liberal society based on strict constraints and norms.

Economics, George Mason University

https://doi.org/10.1017/S0265052510000270 Published online by Cambridge University Press