



Harming and Allowing Harm

Author(s): by David McCarthy

Source: *Ethics*, Vol. 110, No. 4 (July 2000), pp. 749-779

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/233372>

Accessed: 09/07/2014 09:17

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

<http://www.jstor.org>

Harming and Allowing Harm*

David McCarthy

One of the most puzzling structural features of commonsense morality is the presence of asymmetrical constraints on harming and allowing harm or for short, just asymmetrical constraints. By that I mean that the constraints on harming are strong and that the constraints on allowing harm are weak. Very roughly, commonsense morality says that it is typically impermissible for an agent to harm someone else even if the agent would thereby prevent herself or a bystander from suffering a greater harm. But commonsense morality also says that it is typically permissible for an agent to allow someone else to suffer a harm even if the agent thereby only avoids herself or a bystander suffering a smaller harm.

This feature of commonsense morality has come under a tremendous amount of criticism, and some of the most powerful writing in moral theory has concluded that asymmetrical constraints appear to be indefensible.¹ I want to argue that asymmetrical constraints are defensible, and defensible, moreover, on grounds that their critics could accept.

COMMON GROUND

The first ground the critics could accept is the idea, to put it very roughly, that the claim that such and such a morality is correct must ultimately be based on some kind of appeal to what makes people's lives go best.² This

* I presented an earlier version of this at the University of Melbourne and at the Australasian Philosophy Conference in 1999 and benefited from the comments I received, particularly from Michael Smith. I am especially grateful to Frances Kamm for written comments on a still earlier version. Kamm's written comments were so detailed and extensive that it would take me at least another article to try to respond adequately, and I regret that space constraints have pushed many of the issues she raised to the side. I am also grateful to Judith Thomson for related correspondence and grateful to the referees and editors of *Ethics* for extensive comments and helpful suggestions.

1. See, e.g., Jonathan Bennett, *The Act Itself* (Oxford: Clarendon Press, 1995); Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982); Shelly Kagan, *The Limits of Morality* (Oxford: Oxford University Press, 1989).

2. Terminology: I will regard a number of different sets of sentences or propositions about the permissibility of various actions as moralities as long as they have the right kind of subject matter (e.g., some sort of concern with some notion or other of harm and benefit), and I will often speak of features of this morality or that. A different use of terminology would refer to these only as candidate accounts for morality, the term 'morality' only refer-

Ethics 110 (July 2000): 749–779

© 2000 by The University of Chicago. All rights reserved. 0014-1704/2000/11004-0004 \$02.00

idea is, I believe, what makes the attractiveness of utilitarianism so enduring. Never mind its distributive insensitivity and its traditional gloss on what makes people's lives go best. The idea that the correct morality in some way or other best serves our interests is very plausible. It is also what makes asymmetrical constraints so puzzling. If the cost the agent avoided is the same, and the harm the victim suffered is the same, and so on, what does it matter whether the agent harmed the victim or merely allowed the victim to suffer the harm? No one puts it this crudely, but this is surely what lies at the heart of the puzzle. Here I share the view that what are sometimes offered as explanations of asymmetrical constraints are just ways of redescribing the puzzle. Suppose, for example, that we are told that agent-victim relations are worse in the harming case than in the allowing harm case. Choking back the question of what "worse" is supposed to mean in contexts like this, we can still claim that we have not been told how some kind of appeal to what makes people's lives go best makes the one relation worse than the other. Thus from the outset, I will take as my methodological stance the idea that a satisfactory defense of asymmetrical constraints would somehow be based on an appeal to what makes people's lives go best.³

The second ground the critics could accept is the idea that options are defensible. To say that an agent has the option to do such and such is just to say that it is permissible for the agent to do such and such. And the claim that options are defensible is short for the claim that the correct or the most defensible morality gives agents broad options. Commonsense morality gives us broad options, and this is another of its structural features. Add asymmetrical constraints to options, and we get its distinctive structure.

The most powerful criticisms of asymmetrical constraints have been sympathetic to the claim that options are defensible.⁴ I will claim that if options are defensible, then so are asymmetrical constraints and that, moreover, they are defensible by an appeal to what makes people's lives go best. I begin with two well-known rationales for options.

OPTIONS

Scheffler's rationale for options is this.⁵ He claims that states of affairs can be ranked from best to worst by looking at the degree to which each

ring to what, in my terminology, is the true or most defensible morality. This is only a word-saving device, and nothing significant hangs on it.

3. Compare Scanlon: "Rights . . . need to be justified somehow, and how other than by appeal to the human interests their recognition promotes and protects?" See T. M. Scanlon, "Rights, Goals, and Fairness," in *Public and Private Morality*, ed. Stuart Hampshire (Cambridge: Cambridge University Press, 1978), pp. 93–111, p. 93. This is a revised version of an article with the same title in *Erkenntnis* 11 (1977): 81–95.

4. Kagan is an exception to this.

5. See Scheffler.

person's interests are satisfied in each state of affairs and then by applying a distribution-sensitive principle to rank the states of affairs. He also claims that a moral theory should take into account what he says is the independent significance of the personal point of view. One way of doing this, he claims, is via the liberation strategy, which permits each person to attach greater weight to her own interests than they are given in the ranking from best to worst. Scheffler claims that the resulting morality gives us options.⁶ He later denies that there is any comparable rationale for asymmetrical constraints and concludes by endorsing a morality which he claims includes options but not asymmetrical constraints.

Bennett's rationale for options is this.⁷ He argues at length that there is no plausible rationale for asymmetrical constraints, and he then asks what the upshot of this is. One natural proposal starts with common-sense morality and then eliminates the asymmetry by raising the strength of the constraint on allowing harm to be equal to the strength of the constraint on harming. This morality will contain inconsistencies, and the natural way of getting rid of these, he claims, is quantitatively, with the result being a kind of utilitarianism: utilitarianism minus both the distribution insensitivity and the traditional gloss on what makes people's lives go best. Bennett then considers the familiar complaint that such a morality is too demanding and regards this as providing a rationale for a less demanding morality, one which contains options but not asymmetrical constraints. In obvious terminology, like the morality endorsed by Scheffler, this morality gives us options via mild symmetrical constraints: agents are permitted to harm and to allow harm as long as the cost to the victim does not exceed the benefit to the agent by a certain amount or proportion. I now describe an issue relevant to almost any rationale for options.

EXERCISING OPTIONS

Suppose two moralities give us different but similarly broad options: they liberate the personal point of view to a similar degree or are similarly undemanding. If the rationales for options are to liberate the personal point of view or to make morality not too demanding, they do not decide between the two moralities. But because the options the two moralities give us are different, it is possible that through their exercise we would

6. Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986). Chapter 9 defends a structurally similar account of options, but it is complicated by the view that something like asymmetrical constraints can be defended by appeal to something like the principle of double effect. For criticism, see Bennett, chap. 11.

7. See Bennett, chap. 9. The critique of asymmetrical constraints takes up almost the whole book. Much of this critique is taken up with the development of two analyses, in nonmoral terms, of the difference between making something happen and merely allowing it to happen. I feel that I ought to have something intelligent to say about how the present defense of asymmetrical constraints connects with these analyses, but I don't.

all be worse off under, say, the first than the second. This would provide a reason in favor of the second morality.⁸

For example, suppose that we were hypothetical contractors of some sort or other and that for the reasons Scheffler and Bennett offer us, we agree that, whatever other features the morality we are going to agree to has, it will give us broad options. But then we notice that although the two different moralities give us similarly broad options, we would all be worse off if we were to agree to the first rather than the second. Surely, given a choice between the two, we would not agree to the first.

Similarly, both of the rationales for options involve a distribution-sensitive principle for evaluating distributions of benefits and burdens. Suppose that this principle is the priority view, which gives greater weight to improvements in the position of the worst off.⁹ Because the options the two moralities give us are different, it is possible that through their exercise the worst off would be worse off under, say, the first than the second, without the increase in the position of the better off being great enough, from the point of view of the priority view, to outweigh this. This would also provide a reason in favor of the second morality.

Any account of options has to be able to look at moralities with comparable but different options and ask how these options would likely be exercised. As far as I can see, both Scheffler and Bennett ignore this. I fully agree with Scheffler and Bennett that it looks hopeless to try to defend asymmetrical constraints directly, say by looking for some kind of distinctive negative value that is somehow mysteriously present in cases of harming but not cases of allowing harm (compare the claim that agent-victim relations are worse in the harming case than in the allowing harm case). And I share their skepticism about the possibility of defending a moral distinction between harming and allowing harm within the structure of a morality which does not give us options. But consider the fact that the structure of the moralities they defend and the structure of commonsense morality both give us options: the former via relaxed but symmetrical constraints, the latter in large part via strong constraints on harming and weak constraints on allowing harm. If we adjust “relaxed”

8. Note that this is not simply the claim that, in a natural extension of the terminology of Derek Parfit (*Reasons and Persons* [Oxford: Oxford University Press, 1984]), the first morality is collectively self-defeating. The second morality might be as well. It is at least arguable that the moralities defended by Bennett and Scheffler are collectively self-defeating in the most natural interpretation of that term given their approaches, so it is quite possible that this is simply a cost they are prepared to accept in exchange for a morality which has other virtues (liberating or not too demanding). I do not have the space to discuss this here. In this terminology, the point in the text is that if two moralities share those kinds of virtues, but the first is collectively self-defeating to a greater degree than the second, then that is surely a strong point against the first regardless of whether being collectively self-defeating is a decisive reason to reject a morality.

9. The ‘priority view’ is Parfit’s term. See, e.g., Derek Parfit, “Equality and Priority,” *Ratio* 10 (1997): 202–21.

and “strong” and “weak” so that the two structures give us comparable options (e.g., comparably liberating or comparably undemanding), the rationales Scheffler and Bennett offer for options do not determine which of these two structures is more defensible. It is therefore possible that by looking at how the different options would most likely be exercised, we could find reasons in favor of one structure rather than the other and, therefore, possible that we could thereby provide a rationale for asymmetrical constraints. Taking up a hint due to Thomson,¹⁰ I will argue that we can do just that.

COMMON ENOUGH GROUND

We now have to put flesh on the bones of the idea of assessing different packages of options by looking at how they would most likely be exercised. The difficulty is that any complete account of how this should be done will have to take a position on a number of controversial issues which lie far beyond the scope of this article. But most of these controversies can be bypassed. I begin with a specific proposal for assessing competing packages of options. I believe that this proposal is part of a moral theory which provides a better rationale for options than the rationales offered by Scheffler and Bennett, but I cannot begin to defend this claim here.¹¹ But the explanation of the key terms that follows the proposal will show how the controversies can mostly be put to one side in developing a rationale for asymmetrical constraints. First for the proposal, which I call the choice theory. For each competing morality, first assess how well-off each person would be if it were common knowledge that everyone were acting from that morality, where how well-off someone is understood, in substantial part, in terms of her choice interests. Then rank the moralities by ranking the associated distributions according to the priority view. The correct or most defensible morality is the highest ranked. Now for the key terms.

Perfect compliance: By assuming that everyone is acting from the moralities in question, the choice theory assumes perfect compliance. On some views, like utilitarianism, there is no important structural distinction between contexts of perfect and imperfect compliance with morality: the significance of the behavior of others is only as a variable in determining which action available to an agent would bring about what is said to be the greatest amount of good.¹² On other views, like some forms of contractualism, morality is seen as cooperative endeavor, and the shift from perfect to imperfect compliance involves a significant shift

10. See Judith Thomson, “Critical Study,” review of *The Act Itself*, *Nous* 30 (1996): 545–57.

11. I try to defend this claim in a manuscript currently entitled “Options and Constraints.”

12. But cf. the account of cooperative utilitarianism developed in Donald Regan, *Utilitarianism and Cooperation* (Oxford: Clarendon Press, 1980).

in the structure of morality.¹³ Nevertheless, everyone agrees that if a morality is in some way defective in the case of perfect compliance, that is a strong mark against it.

Common knowledge: On some views, like utilitarianism, common knowledge of perfect compliance is just a particular kind of case of perfect compliance.¹⁴ But because of the applicability of the results in classical game theory, which also makes very strong common knowledge assumptions,¹⁵ there may be methodological advantages to starting off with the common knowledge case. On other views, like some forms of contractualism, common knowledge is seen as, in some sense, an ideal case, and it is conjectured that understanding the ideal will help us to understand the shift in the structure of morality as common knowledge of perfect compliance breaks down.¹⁶ Nevertheless, almost everyone agrees that the common knowledge case is a useful case with which to start.

Acting from: On some views, sometimes expressed in the utilitarian tradition,¹⁷ moralities ought to be assessed by assuming merely that people are acting in accord with them. On other views, sometimes expressed in the contractualist tradition, moralities ought to be assessed by assuming that people are acting from them. The assumption that people are acting from the moralities in question is central to the choice theory's derivation of options. Nevertheless, which side of the dispute one takes here turns out not to be important in the discussion of asymmetrical constraints below.

Choice interests: On some views, as typically expressed in the utilitarian tradition, the morally basic understanding of how well-off people are is in terms of their noninstrumental interests. On other views, often expressed in the contractualist tradition, the morally basic understanding of how well-off people are is in terms of some measure or other of what they are able to do¹⁸ or, as I shall say, a measure of their choice interests. The assumption that choice interests are the correct morally basic measure is central to the choice theory's derivation of options. Nevertheless, what we mostly need here is simply some notion of harm,

13. A recent example of such a view is Liam Murphy, "The Demands of Beneficence," *Philosophy & Public Affairs* 22 (1993): 267–92.

14. But again, cf. Regan.

15. For a discussion, see Brian Skyrms, *The Dynamics of Rational Deliberation* (Cambridge, Mass.: Harvard University Press, 1990).

16. A similar conjecture is expressed in, e.g., John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), pp. 8–9, 244 ff.

17. See, e.g., Parfit, *Reasons and Persons*, chaps. 1 and 2.

18. See, e.g., John Rawls, "Social Unity and Primary Goods," in *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982), pp. 159–85; and T. M. Scanlon, "Value, Desire and the Quality of Life," in *The Quality of Life*, ed. Martha Nussbaum and Amartya Sen (Oxford: Oxford University Press, 1993), pp. 185–200. Compare also the account of capability sets in Amartya Sen, "Well-Being, Agency, and Freedom," *Journal of Philosophy* 82 (1985): 169–221.

which either side of the dispute can supply, so which side of the dispute one takes is not very important in the discussion of asymmetrical constraints below.

Competing moralities: Bennett and Scheffler offer independent arguments for broad options and, on the basis of these arguments, might regard moralities which do not give us broad options as not serious competitors to start with. The choice theory does not reject such moralities for independent reasons but simply argues that moralities with broad options are ranked above moralities with narrow options, so that moralities with narrow options turn out not to be serious competitors.¹⁹ Nevertheless, all defenders of options agree that at some point or other, moralities with narrow options cease to be serious competitors, so we need not fuss over which moralities are counted as competitors at the outset.

At least in the case of common knowledge of perfect compliance, all defenders of options can agree that the procedure given by the choice theory for assessing competing packages of options is approximately correct and that the remaining differences are not that important in the discussion of asymmetrical constraints that follows. Asymmetrical constraints have met with such powerful criticism that, even if we can only defend them in this special context of common knowledge of perfect compliance, we will have done much to answer the critics. I will, however, try to say a little about how this defense extends further.

ALLOWING HARM

To work out how well-off people would be if it were common knowledge that everyone was acting from a particular morality, we need to work out how people would behave under that morality. To do that we need to work out how they could behave and then add a plausible theory of motivation. This and the next two sections look at how people could behave under constraints on harming and allowing harm of various strengths. Later sections add the theory of motivation and evaluate the results.

I begin with a simple definition of a constraint on allowing harm and argue that it enables people to externalize the costs of their activities. I later show that various ways of modifying the definition do not affect the argument.

The simple definition is this. Say that (a morality says that) there is a constraint on allowing harm when (it says that): (i) other things being equal, it is impermissible for one person to allow another person to suffer a harm if and only if the magnitude *H* of the harm sufficiently outweighs the magnitude *C* of the costs of preventing the harm; and (ii) for some harm and some cost, *H* does sufficiently outweigh *C*.

A full account of this definition would be a complicated business. It

19. Very roughly: if someone is acting from a morality with narrow options, there will be little she is able to do; hence, in terms of her choice interests, she will be badly off.

would say what the “other things being equal” clause is meant to cover, but I will simply assume that it is there to cover clashes with other constraints and fend off inconsistencies that would otherwise be generated by dilemmatic cases. It would also give us some guidance on how harms aggregate if the rescuer can prevent more than one harm and on how costs aggregate if there is more than one cost.²⁰ But for simplicity, I will assume that we are dealing only with situations in which either the potential rescuer saves the potential victim at the cost of making either the rescuer or a bystander (but not both) bear a simple cost, with no harm or cost to anyone else, or the potential rescuer does not save the potential victim, in which case the victim suffers a simple harm, with no harm or cost to anyone else.

By varying what “sufficiently outweighs” means, we can get constraints of different strengths. If H never sufficiently outweighs C, clause ii says that there is no constraint on allowing harm. If H sufficiently outweighs C if and only if H is much greater than C, we get a weak constraint on allowing harm (i.e., one that relatively seldom prohibits allowing harm). And if H sufficiently outweighs C if and only if H is greater than C, we get a strong constraint on allowing harm. It is possible to have constraints which are far stronger than that: consider a constraint which would require me to cut off one of my legs to prevent you from losing a finger. But no one believes in such a constraint, and the live debate is in the range from no constraint on allowing harm to a strong constraint. Commonsense morality contains only a weak constraint on allowing harm.

I now claim that any constraint on allowing harm, even a very weak constraint, will enable people to externalize the costs of their activities in the following sense. Suppose that there is a constraint on allowing harm. Suppose that there is some end that I would like to pursue, but pursuing it would leave me before a threat of harm which only you could avert. But I know that the magnitude of this harm would sufficiently outweigh the cost to you of preventing me from suffering it. Thus I can pursue my end, knowing that you will prevent me from suffering the harm. Here we rely on the assumption that it is common knowledge that everyone is acting from morality. The net result is that I will have been successful in pursuing my end and will have avoided its cost by you bearing a different cost.²¹

20. Compare the rich discussion of the Tradeoff Idea in Judith Thomson, *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990), chap. 6. I have benefited from the comments of an anonymous referee here. For simplicity, I ignore the possibility that “sufficiently outweighs” might be interpreted by some more complicated function of H and C.

21. The term ‘externalize’ and its cognates are from economics. Compare Schelling: “‘Externality’ . . . refers to the effects of a firm’s actions, or an agency’s or a person’s, that are beyond and outside—‘external to’—the firm’s accounting or the agency’s purview or

A variation on the well-known expensive tastes example will illustrate this. Suppose that my passion is gourmet food. If there is a constraint on allowing harm, then I can spend all my resources today on gourmet food, so that tomorrow I am left facing a threat of starvation. But the harm associated with starvation sufficiently outweighs the cost to you of giving me some of your food, so I can enjoy the gourmet food but avoid starvation by you bearing a certain cost.

Even a very weak constraint on allowing harm might tip the balance: it is worth being famished for but not facing starvation for gourmet food. So any constraint on allowing harm can in principle be exploited to make someone else bear at least some of the costs of pursuing an end. But in any context in which a constraint on allowing harm of a certain strength can be exploited, a stronger constraint on allowing harm can also be exploited, but not necessarily the other way around. Thus the stronger the constraint, the more exploitable it is, and the greater the degree to which it enables people to externalize the costs of their activities.

I am not yet claiming that there is anything in any way unsatisfactory about a morality which enables people to externalize the costs of their activities. I will argue later that there is something unsatisfactory, but that is a matter of substantive argument for a disputed conclusion. Until then I am simply trying to establish the extent to which constraints on allowing harm (and, in the next section, harming) enable externalization to occur. But if I were to try to argue now that there is something unsatisfactory, it would be reasonable to respond that I have not shown very much. So far, I have only shown that a constraint on allowing harm enables people to externalize the costs of their activities in contexts in which they knowingly put themselves in front of threats of harm. So even if there is a problem with a morality that permits cost externalization, this does not show much. The kinds of cases where it is plausible that there is a constraint on allowing harm, the response continues, are ones in which people do not knowingly put themselves before threats of harm.

the person's interests and concerns, but within the accounting, the purview, or the interest of somebody else" (Thomas Schelling, *Micromotives and Macrobehavior* [New York: Norton, 1978], p. 213, n. 13). As I am using the term, it is slightly misleading to say that people are externalizing the costs of their projects, as the cost they are avoiding is a different and, in all the cases I will discuss, greater cost than the cost others bear by their avoiding it, but the advantages of the image seem to me to justify the slight inaccuracy. Similarly, if there is only one cost of an activity, a constraint on allowing harm may only result in some of that cost being externalized, and if there is more than one cost, one cost may be externalized while another is not. But to save words, whenever I say "externalize" these qualifications are implicit. I used to say that in the kinds of cases that follow, the person who externalizes a cost is free riding on someone else's resources, but this is seriously misleading: 'free riding' as used by almost everyone in the economics literature has a much narrower meaning, typically connected with problems of public goods. Some people regard all talk of externalization as confused; I will discuss this later.

The response misses the fact that a constraint on allowing harm is exploitable in cases where someone only knowingly puts herself before a risk of a threat of harm. To illustrate, suppose a mountain climber knows that there is a 10 percent risk of getting caught on a rock face, where she will either be rescued or die. The cost of rescue is one thousand dollars. Suppose that there is a constraint on allowing harm and that the death of a young woman sufficiently outweighs one thousand dollars. The mountain climber will effectively be able to pursue her project but avoid a 10 percent risk of death by making others bear a 10 percent risk of a one thousand dollar loss. Thus a constraint on allowing harm is exploitable not only in cases where people can knowingly put themselves before threats of harm but also in cases where people can knowingly put themselves before risks of threats of harm. These are vastly more common.

Moreover, a constraint on allowing harm is not only exploitable where people know the exact risks but also in cases in which people only know that there is some risk: the mountain climber knows there is some risk of getting stuck but does not know what it is. And it is exploitable in cases even where the risk is more amorphous than that: Cautious and Carefree live side by side. Carefree spends her money as it comes, living life to the full, without much regard to the future. Cautious puts money aside each month and has rainy day insurance. They regularly chide each other: "It's always prudent to put something aside." "You worry too much!" One day, out of the blue, a rainy day befalls Carefree, and she lies before a threat of harm which sufficiently outweighs the costs of rescue. Thus if there is a constraint on allowing harm, Carefree will have been able to pursue her ends by making Cautious bear some of the costs. We will later see other examples of this sort, but this is enough to show that any constraint on allowing harm enables people to externalize the costs of their activities in a wide range of cases, and the stronger the constraint, the wider the range.

HARMING

Say that there is a constraint on harming when (i*) other things being equal, it is permissible for one person to harm another person if and only if the magnitude C of the cost thereby avoided sufficiently outweighs the magnitude H of the harm, and (ii*) for some harm and some cost, C does not sufficiently outweigh H. The various explanations and qualifications of this definition are analogous to the case of allowing harm and need not be repeated.

By varying what "sufficiently outweighs" means, we can get constraints of different strengths. If C always sufficiently outweighs H, clause ii* says that there is no constraint on harming. If C sufficiently outweighs H if and only if C is greater than H, we get a weak constraint on harming (i.e., one that relatively often permits harming). If C sufficiently outweighs H if and only if C is much greater than H, we get a

strong constraint on harming. And if C never sufficiently outweighs H, we get a maximally strong constraint (for short, a maximal constraint) on harming.²² While the live debate on constraints on allowing harm is limited to about half the possible strengths they can take, some people take seriously the possibility of no constraints on harming: on some models, the state of nature contains no constraints and, hence, no constraint on harming.²³ Commonsense morality contains a strong constraint on harming.

There is a pleasing asymmetry between harming and allowing harm. We just saw that any constraint on allowing harm enables people to externalize the costs of their activities, and the stronger the constraint, the greater the extent to which this is possible. I will now argue that any nonmaximal constraint on harming enables people to externalize the costs of their activities, and the weaker the constraint, the greater the extent to which this is possible.

Say that a constraint on harming and a constraint on allowing harm are duals just in case the interpretation of “sufficiently outweighs” is quantitatively the same in each. Thus a strong constraint on harming and a weak constraint on allowing harm will be duals, as will a maximal constraint on harming and no constraint on allowing harm.

Suppose that pursuing a certain activity puts me in the way of a threat of harm and suppose that I, or someone on my behalf, can extract aid from the bystander at exactly the same cost to the bystander as the bystander preventing me from being harmed. A constraint on allowing harm obliges the bystander to prevent me from suffering the harm at a particular cost if and only if the dual constraint on harming permits me, or someone on my behalf, to make the bystander bear a cost of equal magnitude to prevent me from suffering the harm. Thus, in this case, I can externalize the cost of my activity by making the bystander bear it via a constraint on harming if and only if I can do the same via the dual constraint on allowing harm. Note that unlike the case of allowing harm, we do not here need the common knowledge assumption.

22. I say “maximally strong” rather than “absolute” for reasons similar to the reasons Thomson gives for calling some claim rights “maximally stringent.” See Thomson, *The Realm of Rights*, pp. 168–69, 197–200.

23. Quinn claims that it is incoherent for a morality to have a constraint on allowing harm that is stronger than a constraint on harming. To modify and extend Quinn’s example, suppose I can save A by killing B. I remove the threat to A’s life by initiating a threat of death to B as this morality may seem to require. But if I can remove the threat to B’s life by initiating a threat to A’s life, then the morality may then seem to require me to do that, requiring me to oscillate faster and faster between saving A by threatening B and saving B by threatening A. There are ways out of this incoherence, but since no one believes in this kind of morality, I will avoid the trouble by stipulating that the constraint on harming is always at least as strong as the constraint on allowing harm. See Warren Quinn, “Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing,” *Philosophical Review* 98 (1989): 287–312.

To illustrate, consider again the expensive tastes case where I spent all my resources yesterday on gourmet food and now face starvation. You can improve my situation by giving me food. A strong constraint on allowing harm will oblige you to give me a certain amount of food if and only if the dual constraint on harming, a weak constraint, will permit me to eat that amount of your food or permit a third party to take that amount and give it to me. We already know that any constraint on allowing harm enables people to externalize the costs of their activities; this shows that any nonmaximal constraint on harming also enables people to externalize the costs of their activities. Moreover, just as the extent to which this is possible under a constraint on allowing harm is greater the stronger the constraint, so too the extent to which it is possible under a constraint on harming is greater the weaker the constraint.

Furthermore, this equally extends to cases in which people put themselves before risks of threats of harm. If a constraint on allowing harm obliges people to come to the rescue of the mountain climber, so too the dual constraint on harming will permit someone on the mountain climber's behalf to take someone else's helicopter and come to the rescue or will permit the mountain climber to help herself to the helicopter conveniently parked halfway up the mountain. Thus if a constraint on allowing harm enables the mountain climber to avoid a risk of harm by others bearing a 10 percent risk of a one thousand dollar loss, so too does the dual constraint on harming. Similar remarks apply to the other cases discussed in the previous section. Hence there is a wide range of cases in which a nonmaximal constraint on harming enables people to externalize the costs of their activities, and the weaker the constraint, the wider the range.

INTERNALIZATION ON THE CHEAP?

We have now seen that any constraint on allowing harm enables people to externalize the costs of their activities, and the stronger the constraint, the wider the range of cases in which there can be externalization. Similarly, any nonmaximal constraint on harming enables people to externalize the costs of their activities, and the weaker the constraint, the wider the range of cases. But it would be reasonable to wonder whether this result is a property that arises from the specific way I have defined these constraints and whether some other way of defining them, perhaps with a different or narrower scope, would avoid the result. Since analogous things can be said about harming as about allowing harm, in most of what follows I will only discuss constraints on allowing harm.

One proposal is to restrict the scope of a constraint on allowing harm to situations in which the person before the threat of harm is not at fault for being there. Only if she is not at fault is there any constraint on allowing harm; if she is at fault there is no constraint.

In order to understand this proposal, we need to understand what

is meant by “fault.” The meaning that best seems to fit the way we ordinarily use “fault” is that someone is at fault for being before a threat of harm just in case she is responsible for being there and that she is responsible for being there just in case the harm is her responsibility, and no one else’s. In other words, she is at fault just in case there is no constraint on allowing her to suffer the harm. This proposal may well best capture what is ordinarily meant by “fault,” but it makes the definition of the proposed restriction circular.

The second understanding of “fault” is that someone is at fault for being before a threat of harm just in case she is there as a result of having behaved impermissibly. There are different versions of this understanding, depending on whether and how they insist that the threat of harm and the reason why the behavior was impermissible are closely connected: compare cutting your hand while breaking into the jeweler’s store versus cutting it because you have gotten drunk on the profit. But no matter. Because there is such a wide range of cases in which a constraint on allowing harm enables people to externalize the costs of their activities, for the proposal to prevent externalization on this understanding of fault it will have to say a very wide range of activities is impermissible: it will have to say that it is impermissible for the mountaineer to go rock climbing when the risk is 10 percent, 1 percent, 0.1 percent, and so on. This is independently implausible, and the very narrow range of activities it leaves permissible ill fits with the kinds of moralities under consideration, moralities which give us broad options.

Another proposal is to adopt a counterfactual restriction on the scope of constraints on allowing harm. Call an *exploiter* of a constraint a person who lies before a threat of harm only because she expects that, because of the constraint, she will avoid the threat: had she not so expected, she would not have placed herself before the threat of harm (or the risk of threat of harm). Call a *beneficiary* of a constraint someone who lies before the threat independently of any such expectation: had she not expected to be rescued, she would still have placed herself before the threat of harm (or the risk of threat of harm). Some people think less well of exploiters than of beneficiaries and might propose that the scope of a constraint on allowing harm be limited by having an appropriate counterfactual restriction built in so that exploitation is not possible. This counterfactual restriction has several difficulties, as I discuss later. But for now, I only note that the counterfactual restriction can still result in people externalizing the costs of their activities: consider the mountaineer who will go climbing regardless of any expectation of being rescued and who thereby becomes a beneficiary of the restricted constraint on allowing harm.²⁴

24. Although I will not attempt it here, we will eventually have to ask what happens when we relax the common knowledge assumption. Toward that end, note that the

The lesson of this section is not surprising. The more the scope of a constraint on allowing harm is restricted, the narrower the range of cases in which people can externalize the costs of their activities. But if any constraint on allowing harm has any scope, externalization will be possible, and the broader the scope, or the stronger the constraint, the greater the range of cases in which externalization is possible. Similarly for harming: if any nonmaximal constraint on harming has any scope, externalization will be possible, and the broader the scope, or the weaker the constraint, the greater the range of cases in which externalization is possible.

THE SIGNIFICANCE OF EXTERNALIZATION

What does this show? Two extreme positions may seem tempting. One says that we do not need to say any more to defend asymmetrical constraints; we have already done it. The other says that this is not going to get us anywhere in defending asymmetrical constraints; our approach is hopeless.

The congratulatory position says that it is simply unfair to externalize the costs of one's activities; it is more or less to free ride on others, and that is a paradigm form of unfairness. Other things being equal, a defensible morality will not permit unfairness; hence, other things being equal, it will contain a maximally strong constraint on harming and no constraint on allowing harm.

But there are two worries about this position. First, it looks as if it may simply be smuggling in an intuition about an asymmetry between the constraints on harming and allowing harm. It might be claimed that to externalize a cost is to make someone else bear that cost, which is, more or less, to harm that person. Thus, this response continues, what we have shown is that any constraint on allowing harm (likewise any nonmaximal constraint on harming) provides a somewhat circuitous route by which people are permitted to harm others, more or less. If we simply claim that, by itself, that provides a reason in favor of no constraint on allowing harm (likewise a maximal constraint on harming), it may appear as if we are simply presupposing the claim that we set out to argue for, that the constraints on harming are strong and the constraints on allowing harm are weak.²⁵ Second, one might even suspect that it is an advantage for a morality to enable us to externalize the costs of our activities in the way we have been discussing: if I can convert any large costs

existence of beneficiaries of either a nonmaximal constraint on harming or a constraint on allowing harm does not depend on the common knowledge assumption. Neither does the existence of exploiters depend on it in the harming case; the common knowledge assumption is only used in the allowing harm case, and it would have been possible to have gotten by with a less strong assumption.

25. Thanks to Smith and Kamm for urging the need to address this response.

I face into smaller costs you bear, and you can convert large costs you face into smaller costs I bear, might not we both think that this is a good deal? If we were hypothetical contractors, might not we welcome such a morality?²⁶

The deprecating position says that all talk about externalizing costs is just confused. The home of this claim is in the law and economics literature, and the target a way of doing economics.²⁷ Some legal theorists, however, have claimed that much moral philosophy has inherited this confusion.²⁸

To illustrate, suppose that I am in my attic, clearing out my old books, and I am too lazy to carry them downstairs to the trash. Instead, I simply throw them out of the window down onto the sidewalk. Just as I throw out my hardbound copy of *War and Peace*, poor Bloggs walks underneath the window and gets a sharp knock on the head and a fractured skull. All sides agree, I think, that Bloggs has been harmed, and the unsophisticated among us—I for one—would say that the harm is a cost of my activity, not Bloggs's. But then we are pressed: tell us what it is about my activity (clearing out my attic by throwing my books out of the window) that makes the harm the cost of my activity rather than the cost of Bloggs's activity (walking home). For example, if we try to say something like: "Had I not engaged in my activity, Bloggs would not have suffered a fractured skull," we will immediately be asked why we do not also say: "Had Bloggs not engaged in his activity, he would not have suffered a fractured skull." No matter what kinds of nonmoral concepts (counterfactuals, causation, and so forth) we try to bring to bear on the matter, it is claimed that we will not get a satisfactory general way of distinguishing between my activity and Bloggs's activity which allows us to

26. I have not seen the phenomenon of cost externalization discussed anywhere in the literature on harming and allowing harm, except for some very brief suggestions in Thomson, "Critical Study." Thomson describes a case in which I can make an outcome come about either by making Smith suffer a harm or by making myself suffer the same harm. She claims that imposing the cost on Smith would be taking unfair advantage. But by contrast, she claims that if I have the choice of allowing Jones to suffer a harm or allowing myself to suffer the same harm, it would not be taking unfair advantage to allow Jones to suffer the harm. She cautiously suggests that this is the right kind of direction in which to defend the claim, in the terminology used here, that there is a difference between the constraints on harming and allowing harm, and there is a suggestion that unfair taking advantage of others is a kind of free riding which hypothetical contractors would avoid if they could. But she seems more inclined to regard an appeal to unfairness as the endpoint of explanation, and, as the text indicates, I disagree. Nor does she try to argue that my imposing a small cost on someone else in order to avoid myself bearing a large cost is unfair. Nevertheless, my overall argument strategy should still, I believe, be regarded as an attempt to develop Thomson's suggestions.

27. The classic expression of this claim is Ronald Coase, "The Problem of Social Cost," *Journal of Law and Economics* 1 (1960): 1–30.

28. For a recent expression of this claim, see Stephen Perry, "Libertarianism, Entitlement, and Responsibility," *Philosophy & Public Affairs* 26 (1997): 351–93.

say that Bloggs's fractured skull was a cost of my activity and not a cost of Bloggs's activity. It is therefore said to be just straightforward confusion to claim that Bloggs's fractured skull is a cost of my activity and, hence, confusion to claim that I have externalized a cost of my activity. The only other possibility, therefore, is to use moral concepts to distinguish between the two activities. For example, one might try to claim that, of the two activities involved in Bloggs's suffering a fractured skull, only mine was wrongful and then use that claim to provide a principled way of assigning the cost to my activity rather than to Bloggs's. But then, the objection continues, the approach being developed here is just wrong-headed. It claims that this or that morality enables people to externalize the costs of their activities in this or that way as if there was some non-moral sense of the costs of an activity when in fact there is no such sense.

I register the opinion that (a) the objection that there is no satisfactory way of assigning costs to activities which only use nonmoral concepts, if sound, would apply with more or less equal force to the ways of using moral concepts to assign costs to activities that the objectors propose in their place, but that (b) there are satisfactory ways of assigning costs to activities (more precisely, expected costs) which only use non-moral concepts.²⁹ But these claims involve technicalities that I cannot go into here, but this does not matter. For the purposes of the arguments that follow, I do not, strictly speaking, need to talk about the costs of activities; all I need is to argue that under different moralities, different distributions of benefits and costs would occur. Suppose that I am acting from a morality under which it is permissible for me to throw books out of my window and that if someone is harmed as a result, I am not under a duty to compensate that person. Then I will throw books out of my window whenever it is convenient, and others will suffer fractured skulls, avoid walking on this piece of sidewalk, or pay me not to throw them out. Compared with a morality under which it is not permissible for me to throw things out of my window, I am better off (more liberty, easier book disposal, or extra income), and others are worse off (fractured skulls, less liberty, or less income). If we have ways of evaluating different distributions of benefits and burdens—and this is what the priority view provides us with—we might be able to say whether the distribution of benefits and burdens associated with it being permissible for me to throw books out of my window is more or less favorably evaluated than the distribution associated with it not being permissible. But in none of this story would we, strictly speaking, have to say that particular costs are costs of particular activities. The subsequent discussion will be easier to follow if we continue to talk of the costs of activities in the unsophisticated way, and since

29. I argue this in "Liability, Joint Contribution, and the Behavior of Victims," manuscript.

the lesson of this paragraph is that we could avoid this talk if we had to, there is no reason why we should not help ourselves to the convenience.

Midway between these two extreme positions, I want to argue that by using the procedure given by the choice theory, we can see that given a plausible view about human motivation, there are reasons to reject any morality in which people are able to externalize the costs of their activities. Moreover, the strength of these reasons increases as the degree to which externalization is possible increases. It follows therefore that there are reasons to reject any morality with any constraint on allowing harm and any nonmaximal constraint on harming, and that the strength of these reasons increases as the strength of the constraint on allowing harm increases and the strength of the constraint on harming decreases.

A SIMPLE MODEL

The essentials of the argument will be easier to follow if we use a simple model. I will assume that there is one all-purpose good and that this good is equally divided among all members of society. Furthermore, assume that people have the option to do whatever they like, including disposing of their share of the all-purpose good in any way they like, aside from the limits set by whatever constraints on harming and allowing harm, if any, that are in place. In other words, (a) each person has the option to do whatever she likes, including disposing of her share of the all-purpose good in whatever way she chooses, unless (b) the constraint on allowing harm requires her to exert efforts or to dispose of some of her share of the all-purpose good to prevent someone else from suffering a harm (e.g., death, confinement, an injury, or loss of a proportion of his or her share of the all-purpose good); and (c) she does not have the option to dispose of any of someone else's share of the all-purpose good, unless (d) the constraint on harming permits her to dispose of some of someone else's share to prevent someone (herself included) from suffering a harm. Let us also assume that there are no gross differences in people's abilities to convert the all-purpose good into functionings. It is not the case, for example, that some require twice as much of it to be well nourished as others.³⁰ With these assumptions in place, it looks at least roughly as though there is at least rough equality as measured by people's choice interests under such a system; call the system the *rough model of equality*. But let us look more closely at what would happen if there were some constraint on allowing harm or a nonmaximal constraint on harming, so that people could externalize the costs of their activities in the way already outlined.

We noted earlier that working out how people would exercise their

30. Sen has emphasized the moral significance of variations in efficiency of conversions of all-purpose goods into functionings (see, e.g., Sen).

options requires a theory of motivation. The theory I will rely on is very simple: people care far more about themselves and their own ends and projects than they do about other people and their ends and projects. Therefore, if people can externalize the costs of activities that they would like to pursue, then by and large they will. For example, suppose that there is some activity A I would like to pursue, but pursuing it creates a cost C . I prefer pursuing A to not pursuing it, but not if I have to bear the cost C . In an obvious notation, my preferences are $A > \text{not-}A > A - C$. I am usually indifferent to whether or not you bear C . This follows from our theory of motivation and the fact that if C is a cost on your resources (including your efforts), then I am usually unable to dispose of your resources or make use of your efforts in any other way: see c (and a) above. So if the morality we are acting from enables me to do A but make you bear C , then I will.

The natural interpretation of the procedure used by the choice theory is to assess people's entire life prospects under competing moralities. Under a morality which enables people to externalize the costs of their activities, there will be some externalization and three possible ways that this can happen, or a combination thereof.

Possibility 1: The predictable result is that by and large things balance out, so that each person externalizes onto others costs of comparable magnitude to the magnitude of the costs that others externalize onto her.

Possibility 2: Some people have an identifiable characteristic (strength, skill, or whatever) that makes them better placed to exploit the possibility of externalizing the costs of their activities than others. Call them the Strong. The predictable result is that the Strong will externalize costs of significantly greater magnitude onto the Weak than vice versa.

Possibility 3: Being in a position to externalize the costs of one's activities is largely a matter of luck and being lucky is not an antecedently identifiable characteristic. But nevertheless the predictable result is that the Lucky will externalize costs of significantly greater magnitude onto the Unlucky than vice versa.

Each of these possibilities, I will argue, is in some way a bad result from the point of view of the priority view. (But note here that we are not yet asking whether there is some other feature of moralities with a constraint on allowing harm or a nonmaximal constraint on harming which in some way compensates, from the point of view of the priority view, for the bad result.)

POSSIBILITY 1

Here is a simple model of this possibility. As above, there is some activity A I would like to pursue but not if I have to bear the costs: my preference structure is: $A > \text{not-}A > A - C$. Likewise for you: you would like to

TABLE 1

	You do B	You do not do B
I do A	Third best for each of us	Best for me, worst for you
I do not do A	Worst for me, best for you	Second best for each of us

pursue B but not if you have to bear the costs. But as it happens, you can pursue B and make me bear the costs, and I can pursue A and make you bear the costs. The situation we face is this, with the rankings based on our preferences (see table 1).

No matter what you do, I am better off doing A. Likewise for you. So we are caught in a prisoner’s dilemma, each ending up with what is third best for each of us when we could have ended up with what is second best for each of us.³¹ (And note that we can describe this kind of structure without having to say which costs are costs of which activities.)

It has often been said that we can avoid prisoner’s dilemmas by forming agreements: I agree not to do A in exchange for you not doing B. Then we could avoid the prisoner’s dilemma provided the costs of forming an agreement were not too high (these include information costs, transaction costs, and so on).³² For example, if the costs of the agreement exceed its benefits, then we are each better off under the third-best option than under the second-best one minus the agreement costs, and it will not be in our interests to form such an agreement. But even if the agreement costs are low enough to make it worthwhile to form an agreement (and they will very often not be) but are still non-zero, there is still a way in which the possibility of externalizing the costs of our activities is making us each worse off: it makes us bear the agreement costs. If we were unable to externalize the costs of our activities, I would not do A, you would not do B, and we would not bear the agreement costs.

All of this is well known. But there is a slight problem in interpreting it within choice theory’s procedure for ranking competing moralities. The discussion of the prisoner’s dilemma is couched in terms of preferences, but our procedure uses choice interests as the morally basic measure of how well-off people are. And the procedure focuses on complete lives lived under a morality, whereas the prisoner’s dilemma happens at a particular time. But it seems very plausible that, by and large, the more often you and I face a prisoner’s dilemma, the worse off we will be in terms of our choice interests. For example, if our choice interests are

31. At least in the context of iterated prisoner’s dilemmas, it has been argued that, under a wide range of conditions, cooperation is rational. But it is very implausible that cooperation is rational in every prisoner’s dilemma. That is all I need for the argument that follows, so I will not try to address the large technical literature on this topic.

32. Perhaps better: if “form an agreement” or some such was an available act, then we would not be facing a prisoner’s dilemma to start with.

understood in terms of the size of our capability sets, the prisoner's dilemma structure will tend to create inefficiency in our conversions of all-purpose goods (and efforts) into functionings; hence, we will tend to have smaller capability sets.

A natural observation at this point is that not all scenarios in which I can make you bear the costs of my favored activities, and likewise for you, will generate a prisoner's dilemma. Suppose that my preferences are not $A > \text{not-}A > A - C$, but instead $A > A - C > \text{not-}A$; and likewise for you. I will do A and you will do B, but this time the outcome is second best for each of us, and we are not caught in a prisoner's dilemma.

This suggests that we could prevent a constraint on allowing harm from generating a prisoner's dilemma by suitably restricting the scope of the constraint. The scope restriction, of course, is that there is a constraint on allowing harm only in circumstances when the person lying before the threat of harm (or risk of threat or harm) as a result of performing a particular activity would have preferred to engage in the activity and suffer the harm (or bear the risk of harm) than not to engage in the activity and not suffer the harm. (With minimal rationality assumptions, this means that there is a constraint on allowing harm only when the person before the threat of harm is a beneficiary, not an exploiter.)

There are, however, at least three difficulties with this proposed scope restriction. First, the scope restriction creates huge information costs: here is someone lying before a threat of harm; what was her preference structure? Quick! Our procedure assesses moralities by assuming that people are acting from them, and, at least to a first approximation, a natural way of modeling that is to assume that people have a regulative desire to act in accord with them (possibly conditional upon others so desiring, etc.). But this means that if someone has a choice between doing F, which she knows is permissible, or doing G, which is permissible only if some further condition K obtains, then she will only do G if she first finds out whether K obtains; she will not take the risk that it does not. So if the cost of finding out whether K obtains is high, it will often not be worth it to her to find out. Since the information costs that the scope restriction creates are typically high, the potential rescuer will very often rescue the potential victim because of the restricted constraint even when the restricted constraint does not tell her she has to, so that a prisoner's dilemma structure still arises. Nor does it help much to shift to a scope restriction based on what the evidence available to the potential rescuer is about the potential victim's preference structure. Too often, there will be too little evidence for this to be meaningful, particularly when being before a threat of harm is the result of accepting a risk of harm. Second, the information the scope restriction requires is information about the preference structure of the person lying before the threat of harm. This fits poorly with the way our procedure uses choice interests to measure how well-off people are: many of the best-known

rationales for using choice interests are based in part upon opposition to using preference satisfaction measures.³³ Third, we will see later that there is a different way of avoiding prisoner's dilemmas (and other problems to be described) which has all the advantages of the scope restriction but avoids the huge information costs and is not based on the structure of people's preferences.

Since we cannot avoid prisoner's dilemmas by suitably restricting the scope of a constraint on allowing harm, it follows that, under the rough model of equality sketched at the beginning of this section, one of the things that could result from a constraint on allowing harm is that everyone gets caught in various sorts of prisoner's dilemmas. They need not be as simple as the one sketched: X could dump a cost on Y, Y could dump a cost on Z, and Z on X so that they are all worse off than if none of them had dumped a cost on any of the others without there being any two-person prisoner's dilemma as a substructure. Nevertheless, if the upshot of a morality which enables people to externalize the costs of their activities would be, or would include, possibility 1, then there is a way in which everyone would be worse off under a constraint on allowing harm than under no constraint, which is a bad result from the point of view of the priority view. Moreover, as the strength of the constraint on allowing harm increases, increasing the degree to which externalization is possible, the worse this result would be from the point of view of the priority view.

POSSIBILITY 2

If possibility 2 obtains, the discussion of possibility 1 shows that there is a way in which the Strong would be better off under a constraint on allowing harm than under no constraint on allowing harm and that the Weak would be worse off. Assume for now that in the rough model of equality everyone would be equally well off if there were no constraint on allowing harm. It would then follow that the Weak would be worse off under some constraint on allowing harm than under no constraint on allowing harm.

Is this a bad result from the point of view of the priority view? Not necessarily. If the gain to the Strong is sufficiently greater than the loss to the Weak, the priority view will say that this is a good result.³⁴ But it seems unlikely that this would happen. For simplicity, ignore the problems mentioned above about extrapolating from information about preference satisfaction to information about choice interests. If I have the preference structure $A > \text{not-}A > A - C$, and I do A and thereby make you bear C, then the typical result will be that the loss is greater to you

33. See, e.g., Rawls, "Social Unity and Primary Goods"; Scanlon, "Value, Desire and the Quality of Life."

34. See Parfit, "Equality and Priority."

than the gain to me. On the assumption that everyone would be equally well off if there were no constraint on allowing harm, it follows that not only is the gain to the Strong not sufficiently greater than the loss to the Weak, it is not even greater.

In discussing possibility 1, we examined a scope restriction on the constraint on allowing harm which would rule out prisoner's dilemmas. An analogous scope restriction is possible here; this time, it would say that there is a constraint on allowing harm only if $A > A - C \gg \text{not-A}$, where " \gg " means "sufficiently greater." This would ensure (continuing to ignore the problems about shifting from preference information to information about choice interests) that if possibility 2 obtained, it would be a good result from the point of view of the priority view. But the analogous scope restriction faces analogous problems: it creates high information costs and uses preference information in the context of a moral theory opposed to doing that, and we will see later that there is a different way of avoiding this kind of problem which has all the advantages of the scope restriction but avoids the huge information costs and is not based on the structure of people's preferences.

Let us turn now to the assumption that in the rough model of equality everyone would be equally well off if there were no constraint on allowing harm. If this assumption is false, it is possible that the improved position of the Strong, in comparison with their position under no constraint on allowing harm, and the worsened position of the Weak would result in a more equal distribution, one more favored by the priority view. In other words, the possibility of cost externalization would have some kind of compensating role. But while this is in some sense possible, it is also hard to believe. Consider who the Strong would be. Our examples suggest that they would include those with expensive tastes and high temporal discount rates, and they would also include those who would be more skillful or ruthless at exploiting the possibilities of cost externalization. It is hard to see that, in terms of their choice interests,³⁵ there is some reason to expect such people to be worse off than others if there were no constraint on allowing harm. In fact, we will see later that being skillful is, if anything, a predictor of being better off than others under no constraint on allowing harm. Overall, we can ignore the possibility canvassed in this paragraph and conclude that if the upshot of a morality which enables people to externalize the costs of their activities would be, or would include, possibility 2, there is a way in which an identifiable group of people would be worse off under a constraint on allowing harm than under no constraint without those who

35. Recall that within the familiar rationales for using choice interests as the basis for interpersonal comparisons in the context of moral theory, having expensive tastes is not regarded as relevant to a morally basic measure of how well-off people are. See, e.g., Rawls, "Social Unity and Primary Goods."

would be better off being sufficiently better off, which is a bad result from the point of view of the priority view.

POSSIBILITY 3

The difference between possibilities 2 and 3 is that possibility 2 involves scenarios in which having certain personal characteristics is a relevant predictor. If possibility 3 obtains, the result will be structurally similar to the result of possibility 2. But the difference is that rather than a particular group of people being favored by the existence of a constraint on allowing harm, so that some are prospectively better off under such a morality and some are prospectively worse off, everyone faces a kind of gamble: you might turn out to be a member of the Lucky, but you also might turn out to be a member of the Unlucky. Let us not make knowledge out of ignorance and say that each person faces the same probability of being in each group; rather, this gamble involves great uncertainty. What does our procedure view say about this? First, although the procedure assesses things prospectively, it can simply claim that everyone facing uncertainty does not mean that everyone has equal prospects; rather, their prospects are ill-defined. And since it claims that the most defensible morality is one whose associated distribution has a particular shape, it can claim that since having an ill-defined shape is not having a particular shape, the most defensible morality will not have an ill-defined shape. Second, people are notoriously averse to uncertainty. This is not the place to try to assess this,³⁶ but if there are distinctive respectable costs associated with uncertainty, these will provide an additional reason for the obtaining of possibility 3 to be a bad result from the point of view of the priority view.

In summary, if people can externalize the costs of their activities under a particular morality, a plausible theory of human motivation says that they will, and the upshot will be one or a combination of the three possibilities. This upshot is a bad result from the point of view of the priority view and gives our procedure a reason to select against any constraint on allowing harm and any nonmaximal constraint on harming, with the strength of this reason increasing as the strength of the constraint on allowing harm increases and the strength of the constraint on harming decreases.

It is possible, however, that a constraint on allowing harm or a non-maximal constraint on harming has features which are good from the point of view of the priority view and which would provide our procedure with a reason to select for a constraint on allowing harm and a non-maximal constraint on harming.

36. For rich material on uncertainty, with further references, see Susan Hurley, *Natural Reasons* (New York: Oxford University Press, 1989), pp. 371–82.

LOG CABIN

A hiker is caught in a severe blizzard and will die unless she breaks into a log cabin, burns some of the wood, and eats some of the food.³⁷ The widely accepted judgment about Log Cabin 1 is that it is permissible for the hiker to break in. In Log Cabin 2 you are in your log cabin when you hear a knock at the door. You go to open it, and you see a collapsed hiker. You can bring her in, put her by the fire, and give her some food, or you can leave her outside to perish. I have not seen this variation discussed, but I take the judgment that morality requires you to take the hiker in to be equally compelling.

Nothing in my argument is based on attaching any weight to these judgments; they merely suggest a direction for inquiry. But they are independently defensible for two sorts of reasons, both stemming from the priority view. First, it is a feature of human agency that no matter what our projects are, no matter how strong we are, we can sometimes get into scrapes.³⁸ Even the most intelligent, well-informed, and prudent have only limited rationality, information, and resources, and taking appropriate precautions places demands on those resources. Moreover, the majority of these kinds of circumstances arise from our having taken, or not thought about, various kinds of risks, and there is now plenty of evidence for the fact that humans, even well-trained statisticians, are uniformly poor at thinking about risk. And this kind of vulnerability is a general feature of human agency rather than a vulnerability that is specifically tied to the pursuit of certain sorts of activities, which is not, of course, to deny that some activities are riskier than others.

One way of responding to this kind of vulnerability is for people to make agreements to help each other. We have already discussed a similar sort of response to the prisoner's dilemma discussed earlier, and the limitations on agreements there apply equally here. Agreement making is not costless, and, in the kinds of situations we are concerned with, it is extremely expensive, largely because of information and transaction costs: if you are going on a hike, who might you think of making an agreement with? The cabin owners, the cabin renters, other hikers? And how would you contact them? So if morality requires a prior agreement for it to be permissible to break into the cabin, or to be obligatory to take the hiker in, very often such agreements will not be made, or, if they are, the benefits of the agreement will be reduced by the agreement costs. So

37. This comes from Joel Feinberg, "Voluntary Euthanasia and the Inalienable Right to Life," *Philosophy & Public Affairs* 7 (1978): 93–123. It is discussed extensively by Thomson in a number of papers reprinted in her *Rights, Restitution, and Risk* (Cambridge, Mass.: Harvard University Press, 1986). In particular, see "Self-Defense and Rights," "Some Ruminations on Rights," and "Rights and Compensation."

38. For a convincing argument that this is true even of stoics and Philippa Foot's "strong atheist," see Barbara Herman, "Mutual Aid and Respect for Persons," in her *The Practice of Moral Judgement* (Cambridge, Mass.: Harvard University Press, 1993).

there is a way in which everyone is better off for there being a constraint on allowing harm. Hence, from the point of view of the priority view, there is a different sort of reason in favor of there being some constraint on allowing harm.

Libertarians often claim that there is no constraint on allowing harm and a maximal constraint on harming. One response to this claim is that it simply overlooks the vulnerability of human agency. But suppose defenders of this claim concede up to a point: they are willing to accept a very weak constraint on allowing harm and a slightly less than maximally strong constraint on harming, as long as it is to their advantage. But although such morality may be well suited for those who happen to be relatively good at thinking about risk, are relatively well-informed, enjoy thinking about complicated insurance schemes, and so on, many people are not like this. Those who happen to be less good will tend to do less well, so a further strengthening of the constraint on allowing harm and weakening of the constraint on harming will improve the situation of the worst off (read nonrigidly) and up to a point will be good from the point of view of the priority view (the last two sections showed that it does not matter whether the worst off are prospectively identifiable).

So far, then, we have seen, from the point of view of the priority view, a reason in favor of no constraint on allowing harm and a different kind of reason in favor of a strong constraint. One possibility is that one of the reasons is strong enough to cancel out entirely the force of the other reason, with the winner being either no constraint on allowing harm or a strong constraint on allowing harm. This is somewhat implausible. More plausible is that there is a weak constraint on allowing harm whose strength is such that the two reasons cancel out: any shift in strength closer to or further away from no constraint on allowing harm, and one of the reasons becomes stronger than the other. Analogous reasoning points to a constraint on harming which is neither maximal nor weak but strong. But we can do even better.

HOW COSTS CAN BE INTERNALIZED

Consider harming for a change. The reason in favor of a maximal constraint on harming comes from the fact that any nonmaximal constraint on harming enables people to externalize the costs of their activities. By itself, this reason most favors a morality in which the constraints associated with harming do not enable people to externalize the costs of their activities. But there is more than one way to do that. One way is via a prohibition rule: prohibit people from harming others. Another is via a consent rule: permit people to harm others if and only if the others consent. Each of these rules essentially prevents people from externalizing the costs of their activities by harming others. If there is a maximal constraint on harming, all harming is covered by one or the other of these

rules. But a third way is via a liability rule: permit people to harm others if they can fully compensate their victims for the harms but then require that they fully compensate their victims for the harms. This rule also prevents externalization: although it permits harming, it requires that the costs of the harm find their way back to the person who harmed someone.

In Log Cabin 1 the judgment that it is permissible to break into the log cabin is not all of the story. It is also standard to judge that, other things being equal, the hiker is then required to compensate the cabin owner. In other words, the standard judgment is that this kind of case is governed by the liability rule. I suggest that to a good approximation,³⁹ a morality in which all harming is governed by the appropriate one of the prohibition, consent, and liability rules has all the advantages of a maximal constraint on harming—the fact that it does not enable people to externalize the costs of their activities—and all the advantages mentioned in the previous section of a nonmaximal constraint on harming: compare how the liability rule operates in Log Cabin 1. Thus if the liability rule is in place in a suitable range of cases, we avoid the problems of cost externalization and also the problem that those who end up in serious scrapes, like the hiker in Log Cabin 1, cannot get out of them.

I register the opinion that the liability rule is a powerful but underappreciated device.⁴⁰ There are two main issues. (1) Except in fairly unusual circumstances, the prohibition rule is difficult to motivate. If the general story being developed here is correct, so that almost (see below) all cases of harming are covered either by the consent rule or the liability rule, the main question will be how to determine which rule applies in which circumstances. (2) To keep things simple, I have been trying to talk of risk as little as possible, but I believe that the whole story is better understood in terms of risk. Thus what I really see the story as motivating is a morality with strong constraints on imposing risks of harm and weak

39. Many of the fine-tunings of this claim are fine-tunings of the liability rule or, more precisely, fine-tunings to the account of which of many liability rules applies in which of many different kinds of circumstances. See Thomson, "Remarks on Causation and Liability," *Philosophy & Public Affairs* 13 (1984): 101–33; reprinted in *Rights, Restitution and Risk*; and my "Liability and Risk," *Philosophy & Public Affairs* 25 (1996): 238–63, for a discussion of a range of liability rules and their relative advantages. Note also that sometimes the problems associated with externalization will be minor relative to the transaction costs associated with a liability rule, so sometimes we may simply have to tolerate a small degree of externalization.

40. Its importance has, however, been emphasized by Thomson in a number of papers. See, e.g., "Self-Defense and Rights," "Some Ruminations on Rights," and "Rights and Compensation." Mischief may have resulted from the fact that the debate about harming and allowing harm has often focused on killing and letting die. Since death is usually a noncompensable harm, this may obscure the importance of liability rules to commonsense morality, and that in turn may obscure the extent to which the structure of commonsense morality reflects a concern with cost internalization.

constraints on allowing people to bear risks of harm.⁴¹ I believe that we can solve a number of otherwise intractable problems if we take this approach, including (and it is here we have to pay attention to this problem) the problem mentioned earlier of assigning costs to activities (or, better, the problem of determining liability: for what proportion of the costs of a fractured skull am I liable to Bloggs?); that in the comparison between the consent rule and the liability rule, it is in contexts of risk imposition that the liability rule really comes into its own; and that the structural features of commonsense morality are best described—and, if the story offered here is correct, best defended—in terms of risk imposition. But how the liability rule operates in contexts of risk imposition is, I believe, quite complex, and I cannot say anything useful here.

So far, in this section, I have just been discussing harming. But exactly analogous things can be said about allowing harm. Just as there are three rules which prevent people from being able to externalize the costs of their activities in the context of harming, so there are three rules which serve the analogous purpose in the context of allowing harm. First, there is the dual of the prohibition rule: it is permissible to allow someone to suffer harm. Second, there is the dual of the consent rule: other things being equal, one person is obliged to prevent another from suffering a harm if and only if the first person promised to prevent it. If there is no constraint on allowing harm, all allowing harm is covered by one or the other of these rules. But a third rule is the dual of the liability rule: it is obligatory to prevent someone from suffering a harm, but the person rescued is then obliged to compensate the rescuer for the costs of rescue. Although this rule sometimes requires people to prevent harm, it requires that the costs of the rescue find their way back to the person who was rescued. And again, a morality in which all allowing harm is governed by the appropriate dual of the prohibition, consent, and liability rules has the advantages of there being no constraint on allowing harm—the fact that it does not enable people to externalize the costs of their activities—and the advantages mentioned in the previous section of some constraint on allowing harm: compare how the dual of the liability rule operates in *Log Cabin 2*.⁴² Thus if the dual of the liability rule is in place in a suitable range of cases, we avoid the problems

41. I have tried to outline some of the advantages of thinking about constraints in terms of the imposition of risks of harm rather than the causing of harm in my “Rights, Explanation, and Risks,” *Ethics* 107 (1997): 205–25, “Liability and Risk,” “Actions, Beliefs, and Consequences,” *Philosophical Studies* 90 (1998): 57–77, and “Liability, Joint Contribution, and the Behavior of Victims.” For arguments against this approach, see Thomson, *The Realm of Rights*, chap. 9, esp. pp. 242–48, and “Imposing Risks,” in *Rights, Restitution, and Risk*.

42. My sense is that there is more sympathy for the liability rule than the dual liability rule. I can see no principled reason for this, so the defense of the dual liability rule may be a minor challenge to commonsense morality.

of cost externalization and also the problem that those who end up in serious scrapes, like the hiker in Log Cabin 2, could stay stuck in them.

Are we missing anything? Perhaps we are overlooking restrictions of options.⁴³ If cost externalization is possible, people will take a degree of trouble to avoid others externalizing costs onto them. But this simply means that one kind of cost (bearing the costs of the activities of others) gets converted into another kind of cost (defense costs, opportunity costs) borne by the same people. The greater the degree to which externalization is possible, the greater these costs will be, so considerations of defense costs function in the same way as externalized costs and do not alter the structure of the story. In particular, the liability and dual liability rules get rid of the need for defense costs, at least assuming common knowledge of perfect compliance.

Perhaps we are overlooking restrictions of options. If there is a maximal constraint on harming, our options will be narrower than if there is a nonmaximal constraint on harming, and our choice interests will be correspondingly reduced. This is correct but there are two ways in which the significance of the point is limited. Here we need to talk about risk. Restrictions on our options to make others bear near certain risks of harm have little impact on what we are able to do, especially as our account of a constraint on harming is silent on dilemmatic cases, self-defense, and so on. Restrictions on our options to make others bear low-level risks of harm would have much more impact on what we are able to do, as many of our activities impose small risks of harm on others. But this is the kind of area in which a liability rule will generally apply, so that we have the option to impose small risks of harm on others but, very roughly, are then obliged to compensate the risk bearers. Such a liability rule would be onerous if it actually obliged us to do something each time we imposed a risk. But, in many cases, a better liability rule for governing such cases is one whereby, other things being equal, the risk imposer is obliged to compensate the risk bearer for the harm if and only if the risk bearer suffers the harm.⁴⁴ Where the probability of the harm is small, such a rule will relatively seldom oblige the risk imposer to actually do anything. So although the fact that any constraint on harming involves a restriction of options creates a reason against any constraint on harming, the strength of this reason is weak even in the case of a strong constraint on harming and will have little effect on the structure of our account.⁴⁵

43. The concerns of this and the next paragraph were raised by an anonymous reviewer.

44. See my "Liability and Risk" and "Rights, Explanation, and Risks."

45. It becomes more significant, I believe, in working out the two main issues raised in this section about consent and liability rules.

LIMITS TO THE ARGUMENT

There are two kinds of scenarios to which the argument for asymmetrical constraints does not apply. The rough model of equality assumed that people start out with equal shares of the all-purpose good and asked what would happen under constraints on harming and allowing harm of varying strengths. It concluded that if people are able to externalize the costs of their activities, then either everyone will be worse off or some will be better off and others worse off (or a combination thereof). But if people had very unequal shares of the all-purpose good to start with ("had" here may have to be understood in terms of de facto control or legal entitlement), a shift in resources away from the rich and toward the poor would only improve things from the point of view of the priority view. In particular, nothing in the arguments for asymmetrical constraints supports either of the following claims. (1) The rich are not making the poor suffer but only allowing them to suffer, and since the constraints on allowing harm are weak, the rich are not behaving wrongly. (2) It would be wrong for the poor to take from the rich because the constraints on harming are strong. In his attack on asymmetrical constraints, Bennett acknowledges the disturbing conclusion that, once we reject asymmetrical constraints, it is no longer very clear what is shielding us, the rich, from the very strong demands of the poor.⁴⁶ Although I have argued that asymmetrical constraints are defensible, nothing in this defense shields us from that disturbing conclusion. This does not, of course, entail that nothing else shields us either.

The same point writ small applies to the following kind of case. Suppose that there is a rare disease which is equally likely to strike anybody and that nothing anyone can do would alter his or her risk of being struck by the disease. The disease will quickly kill the person whom it strikes unless the nearest person cuts off one of his or her legs to make an antidote. Since nothing anyone can do would alter his or her risk of being struck by the disease, it is not the case that pursuing any activity increases one's risk of being struck by the disease. A constraint on allowing death in this case therefore does not enable anyone to externalize the cost of any of her activities, and nothing in our defense of asymmetrical constraints will provide any support for the claim that, when someone is struck by the disease, the nearest person is not obliged to cut off one of his or her legs. In fact, the choice theory seems to select a morality which would so oblige the person: isn't it a good deal for everyone to exchange a risk of death for a risk of loss of leg?⁴⁷ But, Kamm objects, we

46. See Bennett, chap. 9.

47. Compare the account of the Trolley problem in Thomson, *The Realm of Rights*, chap. 7.

are certainly not so obliged,⁴⁸ so the account of asymmetrical constraints developed here is either flawed or incomplete.

I say that it is incomplete. In anything remotely resembling the world we live in, I simply would not cut off one of my legs, and I think this is true of almost everyone. So if the person just struck by the disease lying before me were in my shoes, there is overwhelming evidence that she would not cut off one of her legs to save me, and that surely means that she has no complaint against me when I do not cut off one of my legs to save her. And that, I suggest, is enough to defend the claim that morality does not require me to cut off one of my legs to save her. But we are now entering the part of moral theory concerned with breakdowns in (common knowledge of) perfect compliance. I suggest that these kinds of considerations of reciprocity are central to the way the choice theory operates in this area. This lies far outside the scope of this article, but I have said enough, I believe, to establish at least a temporary truce between Kamm's objection and my account of asymmetrical constraints.

Finally, consider a different case. Suppose that, in the rough model of equality, someone spends his entire share of all-purpose resources on drink, is left destitute, and will starve to death unless we help him. We set him up on his feet, and without him paying us back, he does it again and again. Are we permitted simply to let him die? A pure version of the choice theory would say that we were, but that is a hard saying. One way of motivating the choice theory's use of choice interests is to see this as reflecting, in phrasing similar to Rawls's, the fact that we have the capacity to form, revise, and pursue a conception of the good in the light of reasonable expectations about the means available to us. This is a fairly exalted view of our rationality, and if this was all the theory had to say about our rationality, a pure version of the choice theory would result. But our rationality is also limited and vulnerable in many ways, and while all of us are limited and vulnerable in some ways, some are more so than others. Breakdowns in practical rationality are just as threatening to our ongoing rational agency as getting stuck in blizzards, and a version of the choice theory that recognizes this will generate obligations to rescue in such cases just as in Log Cabin 2, and often, I believe, subsequent obligations by the rescuee to pay for the costs of rescue (we all know the risks of drink). And in cases like the one described above in which the breakdown in practical rationality seems irredeemable? Perhaps we hope against hope that the breakdown is not irredeemable, and if the choice theory puts enough weight on the maintenance of the conditions of ef-

48. This objection to a related position was made in Kamm, "Non-consequentialism, the Person as End-in-Itself, and the Significance of Status," *Philosophy & Public Affairs* 21 (1992): 354–89. I am grateful to Kamm for reminding me of the need to respond to this objection.

fective rational agency, rather than the existence of richer and more interesting conditions for it to operate in, it will tell us to rescue the drunkard again and again. Or if the breakdown is truly irredeemable, perhaps the concerns of a less pure version of the choice theory will become more utilitarian and tell us to rescue again and again. Or, most likely, some combination of the two.

CONCLUSION

I have claimed that the familiar rationales for options do not determine which options we have. Since two moralities could give us comparable but different options, it becomes relevant to ask how we would most likely exercise each package of options. If, for example, we would all be worse off under the first package than under the second, the second package is more defensible. At least in the ideal case of common knowledge of perfect compliance, I have used a procedure whose outline otherwise different approaches to options can agree on to claim that to a good approximation, the most defensible morality with broad options includes strong constraints on harming and weak constraints on allowing harm. Or, slightly more precisely, it says that, other things being equal, all harming is governed by the appropriate one of the consent, liability, or prohibition rules and that all allowing harm is governed by the appropriate one of the duals of those rules. But no doubt this is still only an approximation.

The dynamical system generated by a large number of agents, each forming and pursuing their own ends given common knowledge of perfect compliance with a morality that includes broad options, is extremely complex, and the problem of trying to work out how such a system would behave is extremely difficult. My attack on this problem has been crude, and no doubt I have overlooked many more aspects of this problem than the many only space rather than ignorance has prevented me from discussing. And I have had little to say about the extent to which these results hold in contexts of breakdowns of common knowledge of perfect compliance, which raises distinctive and difficult problems.⁴⁹ Nevertheless, I hope I have said enough to make plausible the claim that there is a lot more to be said on behalf of asymmetrical constraints than their critics have supposed.

At some point it would be useful to develop more precisely defined models of the kinds of complex dynamical systems in question so that one could study their behavior more systematically, but the value of such efforts is, I believe, highly dependent on the extent to which the models in question are first discursively motivated. If this account has seemed overly crude, it may perhaps still serve some purpose in that direction.

49. For an entry into some of the problems in working out how systems with breakdowns in common knowledge would behave, see Skyrms, chap. 6.