Can Contractualism Save Us From Aggregation?

Barbara H. Fried
Stanford Law School

Research Paper No. 1781092

Stanford Law School
Crown Quadrangle
Stanford, California 94305

Barbara H. Fried, “Can Contractualism Save Us From Aggregation?”

ABSTRACT

This paper examines the efforts of contractualists to develop an alternative to aggregation to govern our duty not to harm (duty to rescue) others. I conclude that many of the moral principles articulated in the literature seem to reduce to aggregation by a different name. Those that do not are viable only as long as they are limited to a handful of oddball cases at the margins of social life. If extended to run-of-the-mill conduct that accounts for virtually all unintended (in the sense of undesired) harm to others— noncriminal activities that impose some risk of harm on others-- they would rule out all action. Moreover, because such conduct poses an irreducible conflict between freedom of action and freedom from expected harm, it can be regulated only by principles that accept the necessity of making precisely the sorts of interpersonal tradeoffs that contractualism is foundationally committed to reject: tradeoffs in which the numbers count, such that a risk of serious harm to one person can be justified by small benefits to the many.

Can Contractualism Save Us From Aggregation?

- *Barbara H. Fried*[*]

One of the chief motives for turning to contractualism and other forms of Kantian constructivism has been to find a way to place principled limits on the scope of aggregative reasoning.[1] By "aggregative reasoning," I mean arguments that rank-order alternative principles for action at least in part based on their aggregate expected benefits (costs), summed across all potentially affected individuals.[2]

In traditional consent-based versions of contractualism, constraints on aggregation are (in principle) built into the requirement of unanimous (hypothetical) agreement, which precludes summing of costs and benefits across individuals unless all of the affected individuals consent to it. In justification-based models, which judge the reasonableness of a principle by the reasonableness of the reasons offered in support of it, such constraints are (in principle) built into the sorts of reasons that do and do not count. In Thomas Scanlon's version of

---

[*] William W. and Gertrude H. Saunders Professor of Law, Stanford University. I would like to thank the Center for Advanced Studies in the Behavioral Sciences (Stanford, CA), where this article was written. I have profited from discussions of this project with a number of people, including Joe Bankman, Josh Cohen, Barbara Herman, Mark Kelman, Rahul Kumar, Arthur Ripstein, Debra Satz and Philip Pettit. Special thanks go to Leif Wenar for helping me sort out the various strains of contractualist thought. Some of the aforementioned disagree with almost everything written here, and I thank them for their willingness to engage seriously with the project nonetheless.

[1] Scanlon (1982, pp. 103, 267); Scanlon (1975, p. 655).

[2] What we count as benefits and costs and how we measure them is itself a huge issue under any system that requires us to compare well-being across individuals. For present purposes, I think it is sufficient to say that the meaning adopted in the contractualist literature is roughly the meaning adopted by aggregative welfarists: costs and benefits reflect individuals' own subjective evaluations of different outcomes, laundered where deemed appropriate to eliminate morally offensive preferences, or idiosyncratic preferences that are thought not to have a legitimate claim on collective resources.

justification-based contractualism, for example, the only complaints that count for or against a proposed principle are those held by some real or hypothetical individual (the so-called "individual reasons restriction"), and the relative gravity of different individuals' complaints is judged through pair-wise comparisons, without regard to the numbers of others who hold similar complaints. As one commentator put it, "This kind of contractualism…insist[s] that … [h]arm to you cannot be straightforwardly compensated by benefits to me."[3]

Given that contractualism has defined itself in opposition to aggregation, its success must stand or fall on whether it has succeeded in constructing a morally plausible alternative to aggregation-- or, in Joseph Raz's words, it must stand or fall on "the ways in which numbers do and do not determine the outcome of moral conflicts".[4]

In this article, I examine contractualists' efforts to come up with an alternative to aggregation, in the context of duty that has gotten the lion's share of the attention in contractualist and other Kantian arguments about what we owe each other over the past four decades: the duty not to harm others/duty to rescue them from harm.[5] I argue that those efforts have failed. Some of the principles proposed appear to be aggregation under a different name; at the very least, it remains to be seen whether they can be cashed out in a manner that is meaningfully distinct from standard aggregative techniques. The rest *are* clearly distinct from

---

[3] Lenman (2008, p. 100).
[4] Raz (2003, p. 360),
[5] Scanlon succinctly states the moral importance Kantians attach to physical harm to others: "[T]he idea that each person has a special claim to and authority over his or her own life and body" is "one that no one could reasonably reject." (2000, p. 313).

aggregation, but they are viable only as long as they are applied only to a handful of cases at the margins of human interactions. They cannot be scaled up to resolve the more general problem of harm to others. Moreover, although I do not explore this claim at length here, I suggest that the effort to come up with nonaggregative principles *must* fail-- that the problem contractualists are trying to solve by its nature can be solved *only* with the sorts of interpersonal tradeoffs that the contractualist enterprise is committed to rejecting.

A few clarifications of the scope of the argument offered here.

First, this inquiry is limited to the problem of harm to others. I offer no views on the viability of nonaggregative principles in other contexts, although the analysis offered here suggests some of the factors that are likely to make such an effort more or less successful in different domains.

Second, I am interested in whether contractualists have come up with a *substantive* decision rule that imposes meaningful constraints on interpersonal aggregation. Even if they have failed to do so, they may nonetheless have offered a very different rationale from conventional welfarism for getting to roughly the same conclusion (available options are ranked by their expected aggregate net benefits to society). I do not mean to suggest that difference is unimportant. It is simply not the difference I wish to explore here.

Third, for these purposes, I accept the moral plausibility of a mixed theory—one that (in Stephen Perry's formulation) holds not that *all* risks cannot be run, but

that "certain risks cannot be run, so that consequential justification must take place within permissible deontological bounds."[6] The questions I'm interested in are: (i) what are the permissible bounds that contractualists are proposing? (ii) do they follow from contractualist commitments? (iii) are they viable, in the minimal sense that they can be operationalized? and (iv) when operationalized, do they produce morally plausible results?

Fourth, I limit my focus to conduct that is socially productive but imposes some risk of harm on others. Thus, I set to the side the sort of potentially harmful conduct that would be criminalized in most societies. The reason for this is simple. Precisely because such conduct generates no social benefits (or at least none we are willing to count), prohibiting it outright entails no social costs and hence requires no interpersonal tradeoffs (or at least none we care about).[7] In contrast, the forms of conduct I am concerned with here-- mostly accidentally harming others, secondarily the deliberately 'harming some to save others' hypotheticals typified by trolley problems-- all involve difficult interpersonal tradeoffs that drive many to view aggregation as an attractive or inevitable solution. If contractualists cannot provide an alternative to aggregation in these cases, then their disagreements with aggregationists are trivial at best.

Fifth, I am concerned only with the permissibility of *acting in a way that imposes (a risk of) bodily harm on others.* I do not address the ultimate victim's

---

[6] Perry (2001, p. 78).

[7] This is an overstatement. All rules impose interpersonal tradeoffs, by virtue of the fact that they will always be under- or over-inclusive of the sorts of conduct we wish to prohibit. Moving from the moral to the legal sphere, they will also be under- and over-inclusive as enforced, and enforcement imposes costs on all of us to generate benefits for only a few. I set to the side these and other tradeoffs that are endemic to all moral or legal norms.

right to compensation or some other form of redress from the party who caused the harm. How we respond, ex post, to the harm we have caused is an independently important question that contractualist principles may well shed light on. But it is a different question. To the extent contractualists mean to address the ex post compensation question rather than ex ante standards for conduct, this inquiry is simply orthogonal to their concerns.

Finally, all of the approaches I consider start with the assumption that what is wrong with engaging in risky conduct is that *harm* may result to others. A significant strain of recent Kantian-inflected contractualist scholarship rejects that assumption, locating the wrong of risky conduct instead in intangible aspects of the conduct itself: that, by engaging in risky conduct, we fail to treat others as inviolable, or fail to respect all persons equally.[8] Because space constraints preclude me from giving the argument the attention it deserves here, I reserve consideration for another day. For now, I note only that the central challenge facing this approach, in my view, is to show that the respect we owe others is not cashed out into a qualified duty not to harm them, thereby taking on all the difficulties explored here, in two steps rather than one.

Even with these limitations, the relevant philosophical literature is immense. In the space of one article, I could not possibly do justice to all the pieces of any one contractualist's argument, let alone the many differences between different accounts, and I have not tried to do that here. Instead, I have tried to inventory the various strategies that contractualists have tried for setting principled limits on

[8] Notable contributions to this literature include Ripstein, (2006, 2009); Kamm (2007).

aggregation, and assess whether they can accomplish what contractualists are hoping to accomplish with them. Scanlon's argument in What We Owe to Each Other supplies a disproportionate share of the illustrations, because Scanlon has offered the most comprehensive account of contractualist reasoning to date. But this paper is not about Scanlon's or any other contractualist's argument in particular. It is about the resources available to contractualism more generally to construct a viable alternative to aggregation.

This approach necessitates giving at best a fragmented view of any one author's argument, many of which combine a number of strategies. For my purposes here, this would be problematic if the sum is greater than its parts—that is, if multiple strategies used in tandem can solve problems that none of them alone can solve. I have tried to be alive to this possibility, and to explore the interactions among different strategies at least to the extent that contractualists have themselves relied on those interactions in formulating hybrid approaches. While I have undoubtedly overlooked some key moves here and given short shrift to others, my hope is that looking at the problem from a different angle will put some aspects of it in a new (and helpful) light.

1. The challenge for contractualists: finding a middle way between ex ante and ex post perspectives on harm.

As is widely recognized in the social contractarian literature on deriving the just state, the terms of political cooperation that people will (hypothetically) agree to depends on what they know about their situation and their available options at

the time their agreement is sought.[9] Most of the normative work of defending any given outcome of a social contractarian thought experiment lies in defending the epistemological point of view (POV) the author has constructed.

While the epistemological POV of the parties has gotten a lot less attention in contractualist arguments about personal morality, it plays an equally pivotal role in that context. In real life, many forms of knowledge potentially affect the principles any individual would reasonably accept or reject to govern potentially harmfully conduct: the general risks associated with acting pursuant to a given principle, aspects of the individual's own situation that might put her at greater objective risk than others, and her own subjective preferences (over possible outcomes, about risk-taking, etc.). All of this information together gives each individual a rough idea of her own *expected* outcome under a proposed principle. Henceforth, I will refer to decisions made from this epistemological standpoint as the "ex ante POV."

Contemporary contractualists' commitment (contra Rawls and Harsanyi) to individual reasons formulated by differentiated, situated selves leads them, plausibly, to endow representative individuals with all these forms of knowledge.[10] In Rahul Kumar's words, "Valid moral principles are principles that must be justifiable to each person (provided she is appropriately motivated), from her own point of view, *with no informational restrictions*."[11]

---

[9] See, e.g., Rawls (1999, p. 16).

[10] Rawls and Harsanyi would allow individuals to know how they are expected to fare under a given principle. But since they have no individualized knowledge about themselves, individual expectations converge with aggregate expectations.

[11] Kumar (1999, p. 295) .

Scanlonians, however, have typically endowed each representative individual with one other piece of information: how she will *actually* fare under a proposed principle. For the balance of the paper, I will refer to decisions made from the epistemological perspective that endows each representative individual with knowledge of how things will turn out *for her* under a proposed principle as the "ex post POV." This (epistemological) ex post POV is should not be confused with the temporal position such individuals occupy, relative to the adoption and deployment of a principle, when their reasons for or against accepting such a principle are assessed. If individuals just 'happen' to know the consequences to them of adopting a given principle before they must choose whether to accept it or not-- the typical factual posture of hypotheticals in the nonconsequentialist literature on harm to others—then individuals will 'naturally' have access to an ex post POV from a temporally ex ante position. In real life, however, people rarely (never??) know the ex post consequences of adopting a principle before they must commit themselves one way or the other. To the extent contractarians incorporate such uncertainty into their hypothetical choice scenarios, the only way to endow someone in a temporally ex ante position with an ex post POV is to allow her to 'peek ahead' in time to see how things will in fact turn out if the principle is adopted. An ex ante POV should be understood to refer interchangeably to these two ways of acquiring knowledge of ex post consequences of a principle before having to choose whether to accept it.

In most situations, knowing how things will turn out for one's self under a given principle swamps in importance all the individualized information that might be available from an ex ante POV. If we ask someone how Taurek's dilemma (whether to save five people on one rock or one on another) should be handled at a time when she has no reason to think her odds of being in the group of five are anything but five times greater than her odds of being in the group of one, we will get one answer.[12] If we ask her after she is identified as the unlucky one, we will almost certainly get another.

The basic problem facing contractualists is that adopting an ex ante POV proves too little and adopting an ex post POV proves too much, for reasons familiar from the social contractarian literature.[13]

If representative persons are imagined to choose general principles from an ex ante POV (knowing herself and her general position in life but ignorant of the ex post consequences to *her in particular* of agreeing to a given principle), each will prefer the principles that optimize her expected subjective value, given her ex ante preferences, circumstances, etc. Assuming a plausible range of risk aversion and altruism, a normal distribution of preferences about outcomes and evenly distributed upside and downside risks, the principles that optimize a given individual's expected position will be roughly equivalent to the principles that

---

[12] Taurek (1977).

[13] This comment pertains only to the decision whether a given risky act is permissible to undertake, not to the question of whether anyone ultimately harmed by the conduct might have a right to compensation from individuals or from the state. Fairness concerns may well have a lot to say on the latter issue—whether posed from a contractualist perspective (e.g., George Fletcher's argument) or some other nonconsequentialist perspective (e.g., the arguments of Jules Coleman, John Goldberg, Benjamin Zipursky, Ernest Weinrib et al. in favor of a corrective justice approach to tort law).

optimize aggregate wellbeing. (To put it another way, social choice converges with individual choice under conditions of uncertainty.) Thus, we end up with an aggregative solution by a different route. There are cases in which one or more of these assumptions do not hold—for example, cases in which the upside and downside risks are known *not* to be evenly distributed ex ante.[14] But at least in the context of bodily harm, these are the exceptions. Typically, we lack information about individual deviations from the norm and/or the capacity to act on it in a systematic way.

On the other hand, if each representative person is imagined to formulate her complaints from an ex post POV (that is, knowing not just about herself in general but also her ex post fate under each candidate principle), virtually every principle may be reasonably rejected by someone, because any principle that authorizes actions that risk gravely harming at least one person will be rejected by the hypothetical representative person who, peeking ahead, learns she will be the unlucky one. For how many of us would volunteer for what, peeking ahead, we know will lead to our certain death, just so that our fellow citizens might avoid chicken pox, have a new bridge to ease traffic congestion, not have to lose twenty minutes of a World Cup broadcast, or even avoid death themselves?

From the start, Scanlon staked his version of contractualism on the moral imperative of the ex post POV. To quote his early version of the argument for an ex post POV:

---

[14] For discussion of these cases, see section 2.2 infra.

> Suppose that A is a principle which it would be rational for a self-interested chooser with an equal chance of being in anyone's position to select. Does it follow that no one could reasonably reject A? It seems evident that this does not follow. Suppose that the situation of those who would fare worst under A, call them the Losers, is extremely bad, and that there is an alternative to A, call it E, under which no one's situation would be nearly as bad as this. Prima facie, the losers would seem to have a reasonable ground for complaint against A.[15]

In What We Owe To Each Other, Scanlon restates the case for an ex post POV as follows: "In assessing the rejectability of [a] principle" based on burdensomeness, "we can begin… by taking the maximum level of burdensomeness and asking whether that would give a potential agent reason to reject the principle." Where the activity in question has benefits for many people but "involve[s] risk of serious harm to some…. [i]t is obvious what the generic reasons would be for rejecting such a principle from the standpoint of someone *who is seriously injured*."[16]

Most other contractualists writing in a Scanlonian vein have adopted an ex post POV as well for weighting the complaints of the would-be Losers.[17] When combined with the maximin rule adopted in Scanlon's Relative Complaint Model, an

---

[15] Scanlon (1982, pp. 122-23). It has been suggested that Scanlon's substitution of "representative types" and "generic reasons" signals a shift to an ex ante POV. Brand-Ballard (2004). There is some textual support for that reading, but on balance I don't think it can be what Scanlon intended. As Elizabeth Ashford has suggested, to read Scanlon otherwise is to assume that he has jettisoned sub rosa the central commitment of his version of contractarianism: that every candidate principle must be judged based on its actual effects on particular individuals (that is, from an ex post POV). Ashford (2003, pp. 273-302, 277). The better reading, I think, is that the substitution is meant to avoid the impossible informational burdens of having to consider what *every* potentially affected individual *actually* would prefer, as well as to screen out idiosyncratic or objectionable reasons that are not plausible candidates for universal agreement. For textual support for maintaining an ex post POV, see Scanlon (1998, pp. 202-205).

[16] Scanlon (1998, p. 207) (itals added).

[17] See, e.g., Parfit (2003); Parfit (2011). Although most contractualists are formally committed to an ex post POV, they frequently slip into an ex ante one. Parfit, for example, after considering a host of other reasons why one might or might not support killing one to save five, concludes "[W]e have reasons to want such people to believe that [in Taurek-type cases] they ought to save as many lives as possible. We would know that, if our lives were threatened in such an emergency, we would be more likely to be one of the people whose lives were saved." (2011, p. []).

ex post POV produces the following result: Whenever a representative hypothetical person, peeking ahead, discovers she will be the one most seriously harmed by adopting a given principle, she may reasonably reject it, provided that the harm to her will be serious as an absolute matter and that some alternative course of action with a less-bad worst outcome is available. Assuming that her reasonable rejection is dispositive with respect to the principle in question, the result, writ large, is what John Broome described thirty years ago as 'moral gridlock':[18] Virtually every action in the world can reasonably be vetoed by someone. Indeed, avoiding just this form of moral gridlock was one of Rawls's motivations for insisting on his much more stringent version of the ex ante POV built into the veil of ignorance, as it is, more recently, one of Parfit's motivations for defining rationality to include a very high degree of altruism.[19]

Given the dispositive role that knowledge of ex post outcomes potentially plays in contractualist arguments, the assumption that the morally ideal POV from which to judge the complaints of would-be Losers is the ex post POV has gotten surprisingly little attention in the contractualist literature. Most Scanlonian contractualists appear to believe it is logically entailed by their dual commitments to (i) adopt the "individual reasons" restriction, which limits the reasons that count for or against principles to reasons held by a single individual, without regard to the number of other individuals who share those reasons; and to (ii) lift the Rawlsian

---

[18] Broome (1978, pp. 92-94).

[19] Rawls (1971, p. 140); Parfit (2011, pp. xx-xx). On the other hand, Rawls's assumption of extreme risk aversion threatens to produce the same gridlock by a different route. For more recent discussions of the gridlock problem, see Reibetanz (1998, pp. 302-04); Ashford (2003, pp. 298-99); Parfit (2011, p. [ ].)

veil of ignorance with respect to one's own identity, permitting interested parties to know their own preferences, their circumstances, etc., before they choose the general principles they wish everyone to live by.

That view, I believe, is mistaken, and results from conflating the question of *whether* people should be stripped of knowledge they possess at the time their agreement is sought (no, if one is committed to 'thick selves') with the question of *when* agreement should be sought. The latter question is simply orthogonal to either the 'individual reasons' restriction *or* a commitment to thick selves. Jane is Jane when she agrees ex ante that, should we ever face Taurek's dilemma, we should save the five, and she is still Jane when she seeks to renege on that agreement ex post, after learning she is the unlucky one. Which Jane is the right Jane to invoke in assessing 'Jane's' preferences has to be argued for on other grounds.

Having committed to an ex post POV for weighting complaints against a candidate principle, the challenge facing contractualists is to find a way to confine its role so as to avoid moral gridlock, without going all the way to an ex ante POV. In the next section, I consider the principal solutions contractualists have explored. Because I am interested in whether any of these solutions, alone or in concert, can give contractualists a viable middle way, I have classified them by what seem to me to be their moving parts—that is, by what they would instruct us to do differently from straight aggregation (e.g., require greater precaution) or from a strict ex post POV (e.g., override Scanlon's Model when the social costs of abiding by it exceed a specified threshold or when the big loser could have avoided the risk in question). Classifying solutions in this fashion does not always correlate with how

contractualists themselves describe what they are doing, and I have tried to flag the discrepancy in terminology where it may lead to confusion.

2. Strategies for Working Out a Viable Middle Way

In subsection 2.1, I start with the qualification that, in my view, has implicitly carried the laboring oar in most contractualist arguments: Limit the ex post POV to acts that, *as of the time of the agreement*, are 'actually' known to be certain to cause harm to an identified victim; all other cases, in which outcomes are uncertain at the time of agreement, are handled by some version of an ex ante POV.

I then take up a variety of other strategies. The first group (2.2 and 2.3) adheres to a pure ex ante (expected value) POV, but gives a veto to those who face disproportionately greater expected harms if a given principle is adopted. Where relevant, distributionally sensitive ex ante POVs can provide a meaningful and (I believe) desirable alternative to conventional aggregation. But their relevance in the real world is relatively limited. Even where relevant, my guess is that contractualists will balk at applying them when the social costs of doing so are just too great.

The second group starts with a pure ex post POV (2.4, 2.5, 2.6) or a pure ex ante one (2.7), but then qualifies it in a way that pushes it in the other direction. The open question with all of these approaches is whether, when all the dust settles, the principles that result are meaningfully distinct from aggregation.

The third group (2.8 and 2.9) tries to combine ex ante and ex post POVs in some fashion-- in the first case by judging some cases from an ex ante POV and

some from an ex post, in the second by requiring that all cases be considered from both POVs, in a kind of reflective equilibrium. For different reasons, I believe neither approach can generate a viable alternative to aggregation.

The fourth (2.10), proposed in Derek Parfit's recent Tanner lectures, adopts an ex post view, but assumes each of us is altruistic enough, at least in some circumstances, that we could 'rationally' choose to have the numbers count, to our own substantial detriment. This strategy allows Parfit to take the ex post POV at least partway back to full-blown aggregation. But it does so, arguably, at the cost of eviscerating "individual reasons" in favor of impartiality.

2.1. Certain v. Uncertain Harms

One of the most powerful intuitions running through the nonconsequentialist literature is that conduct that is "absolutely certain" to result in harm, typically to identifiable others, is factually and morally distinct from conduct that poses a mere "risk" of harm. As one author put it:

> [W]e are permitted to act in ways which do impose risks, even if we think we cannot act in ways which are certain to lead to death . . . Why do we find one kind of action which brings about the death of five impermissible, where the same outcome we may reliably expect to result from a sequence of actions each of which is only associated with the risk of producing part of that outcome?.... The minimal thing we can say here is to point to a structural difference between the two cases and note that ***we do take this to have moral significance***.[20]

[20] Munoz-Darde (2009, [p. 5] For a different but equally emphatic insistence that acting in a fashion that imposes 'certain harm' and in a fashion that imposes a 'risk of harm' are distinct, both factually and morally, see Parfit (2011, pp. [ ]).

Contractualists typically cash out this intuition by weighting complaints about harms 'certain' to result at the full value of the ex post harm to the hypothetical Losers, and complaints about harms 'uncertain' to result at something less than that (sometimes a conventional expected value measure, sometimes a more complicated balance of considerations that may or may not boil down to expected value).[21]

The importance contractualists attach to the distinction between *absolutely certain* consequences and *possible, probable or (in some cases) even statistically certain* consequences is neatly illustrated by the 'welshing' hypotheticals in the philosophical literature. These are cases in which parties to an initial agreement have an opportunity to rescind their consent *after* they learn they will be the losers under the agreement but *before* the harm is actually imposed.

Frances Kamm's "Ambulance Cases" are illustrative of the genre. Ambulance I requires us to decide whether the town may authorize its ambulances to speed on the way to the hospital, whenever it is the case that doing so will save five patients who would have died had the ambulance not speeded, for every one pedestrian killed as a result of the speeding. Kamm first poses the case from a straightforward ex ante POV: May the town acquire an ambulance and permit it to speed? Yes, says

---

[21] Reibetanz (1998, pp. 301, 308). 'Expected value' is calculated in the same manner as in standard aggregative procedures, equal to the full value of the harm, should it come to pass, discounted by the probability it will befall the complainant. Id. Rahul Kumar endorses the same hybrid approach in (1999, p. 295). He does, however, proceed to qualify the commitment to giving *certain* harms their full ex post weight in a way that, I believe, pretty much undoes it. Your right to take into account any information you actually have, Kumar argues, does not exhaust the analysis. The reasons that you give for your position must also be "impartially justifiable." In the context of certain harms, Kumar cashes out that qualification to mean that you may not appeal to the full benefit (detriment) to you of a given principle; you may appeal only to the amount by which it will reduce the probability that an *average* person will benefit (suffer) from the principle. Id,. at 298. I'm not entirely certain how Kumar intends this to operate, but it appears to weight all benefits and harms on an ex ante basis, producing straight aggregation.

Kamm: "preventing a greater number of equally great losses (deaths) seems sufficient reason to introduce the ambulance." In support of that conclusion, she observes (significantly) that *"[p]eople take risks of death all the time so as to increase their probability of survival, as well as for lesser goods*."[22]

In Ambulance II, Kamm asks whether the town may precommit (through technological constraints or by verbal agreement) *not* to allow the driver to brake to save the pedestrian's life, whenever it is the case that doing so will cause five patients inside the ambulance to lose their lives. Clearly such an agreement would be impermissible, Kamm concludes, because "the death would be imposed deliberately at a time when we know that it can no longer be in the interest of the [pedestrian] to risk it. It was just in her interest ex ante to run the risk of being put in this position later."[23]

Kamm's intuition that we have an unwaivable right under these circumstances to rescind what was, at the time it was given, valid consent, is widely shared. In defense of that intuition, James Lenman argues that a policy that "is not acceptable at every time [to every person] is plausibly acceptable at none." Since, in Ambulance II, there is a moment in time in which one person—the now identified about-to-be victim of the speeding ambulance—can complain that the driver is not taking "reasonable precaution against [his] coming to harm," any prior consent the

---

[22] Kamm (2007, pp. 273, 275) (itals added).
[23] Kamm (2007, p. 273).

victim might have given in Ambulance I to resolve this by aggregation is retroactively rescinded.[24]

If we eliminate the original agreement in Ambulance I (which in Kamm's view is superseded in any event in the case of Ambulance II), the same intuition explains contractualists' widespread agreement about how to handle cases in which the outcomes of available actions are stipulated to be known with certainty ab initio. These include cases in which we must decide whether to inflict certain harm on identified victims to save others (trolley problems and their variants), as well as cases in which we must decide whether to rescue one identified person from certain harm or death at a cost to others of qualitatively lesser goods (lesser harm, money, lost pleasure). Like the rescission cases, in both the 'certain harm' and 'certain rescue' cases there is a moment in time at which it is unambiguously *not* in the interest of the about-to-die person to consent to be sacrificed for the greater good. The only difference from the rescission cases is that that moment occurs the first and only time we seek hypothetical assent.

In such situations, as in the rescission cases, contractualists would permit the person who knows her life is at stake in the decision to withhold consent based on that knowledge. Indeed, Kamm defends the duty of easy rescue in almost identical terms to the right to rescind prior consent:

> [S]uppose that we did argue even for the permissibility of investing in cures for truly minor problems affecting many, such as headaches, rather than in a cure of a fatal disease, on the ground that it is reasonable for each person to take a small risk of being the one who will die in order to have headache

[24] Lenman (2008, p. 117).

> cures at hand for his many, certain-to-occur headaches. This does not imply that here and now we should not save someone from dying from a rare fatal disease, if we could, rather than cure millions of headaches. . . . *It could be wrong to leave him to die on the grounds that it was reasonable ex ante, in order to produce the aspirin for headaches, for each person to take a small risk of dying because no help for him would be available when he fell ill. It is here and now that the irrelevant utilities of headache cures do not aggregate to override saving the life.*[25]

Although Scanlon does not explicitly invoke the distinction between certain and uncertain consequences, I believe it similarly drives his conclusion that our duty to rescue someone from imminent harm is much stronger than our duty not to put him in harm's way to begin with. Suppose, says Scanlon, Jones is trapped under live transmission wires, which are inflicting severe pain on him (although no permanent physical injury). In order to rescue him immediately, we would have to cut off transmission of the World Cup soccer game for twenty minutes, depriving hundreds of millions of viewers of the pleasure they would have received from watching those twenty minutes live. Alternatively, we can leave Jones to suffer for an hour until the game is over. Scanlon concludes that of course we must save Jones now, because it is not permissible to leave someone to suffer an hour of severe pain, merely to give recreational pleasure to others, no matter how numerous those others are.[26]

But suppose we are deciding instead whether to build a new system of transmission towers that will improve the quality of reception for many television viewers. We know that in the course of a project like this, some number of workers

---

[25] Kamm (2007, pp. 36-37) (itals added). For the argument that it is irrational to use different procedures to decide whose wellbeing to invest in ex ante and whom to help here and now, see Brock (1998).

[26] Scanlon (1998, p. 235).

will likely suffer harms as great as Jones's. Scanlon concludes it is permissible to proceed with the project, notwithstanding that it "involve[s] risk of serious harm to others," provided the builders use 'adequate precautions' to reduce that risk. (I take up the meaning of 'adequate precautions' in subsection 6. For now, it suffices to say that it is clearly intended to permit us to impose some risks of serious harm on the would-be Losers that we would not be permitted to impose if the consequences and/or the victim were known with certainty ex ante.) The reason we are permitted to sacrifice the construction workers even as we must save Jones, says Scanlon, is that the construction case involves "failing to prevent *accidental injuries* rather than either intentionally inflicting serious harm on a few people or [as in Jones's case] withholding aid from people who need it."[27] While Scanlon does not explicitly invoke the uncertain/certain harm distinction here, I think this is what he must mean in contrasting 'accidental' v. 'intentional' harms. That is to say, 'accidental' refers to situations in which we act without full knowledge of the consequences ('That was an accident' in the sense of 'Gee, I didn't see that one coming') and 'intentional' refers to the opposite. [28]

If the pure ex post POV is difficult to justify under contractualist principles, this hybrid ex post/ex ante POV seems difficult to justify under any moral principles.

---

[27] Scanlon (1998, p. 236) (itals added). Consent also plays a role in this and other hypotheticals. Scanlon explicitly requires that, for the construction project to be permissible, the construction workers must consent to the risk they are assuming. Consent is surely morally relevant, but its relevance is complicated, and cannot ultimately be separated from our judgment about the reasonableness of the risk to which someone has implicitly or explicitly assented. For further discussion of this point, see infra, section 2.6.

[28] It is not clear what else Scanlon could have in mind here, since in the two scenarios he describes as 'intentional,' he clearly does not mean 'intent' to refer to motive/mens rea; what the two examples have in common is that the actor knows the consequences of acting or not acting with certainty ex ante.

It presupposes distinctions in kind, both factual and moral, between 'certain' and 'uncertain' harms that are doubtful at best. And it holds the "reasonableness" of a given principle hostage to the adventitious factual posture of the hypothetical that must be resolved pursuant to it, producing general principles that are both morally arbitrary and unstable.

2.1.1. Are acts inflicting 'certain' harms and those imposing a risk of harm factually distinct forms of conduct?

The philosophical literature on harm to others has been organized around a set of canonical hypotheticals that, like Kamm's Ambulance II, involve identified victims and choices among harms that are stipulated to be known with certainty ex ante. Think here of the trolley problem oeuvre, Bernard Williams's Jim and the Indians, Scanlon's Jones trapped under the transmission wires, Scanlon's Medical Experiments hypothetical.[29] But in the real world, no conduct, judged ex ante, is absolutely "certain" to harm others. This is true even of harms that are intended (in the strong sense of 'mens rea' or the weak sense of 'foreseeable'). If I point a loaded gun at your head and pull the trigger, I am overwhelmingly likely to kill or seriously injure you, but I am not certain to do so. The gun could misfire; I could have forgotten to load it; the bullet could be deflected by a metal plate in your skull. If I divert the trolley, I may believe I have saved five from certain death at the cost of one life, but I can never be certain ex ante. Perhaps diverting the trolley will cause it to tip over before it reaches the one; perhaps the five would have seen the trolley in time and moved out of the way. A fortiori, what is true of intentionally inflicted

[29] Williams (1973, pp. 97-99); Scanlon (1998, 208-09, 235).

harms is true of accidental harms. In predicting the unintended consequences of conduct, we are always dealing with greater and lesser degrees of uncertainty. Thus, from an ex ante perspective, the problem of harm ***is*** the problem of risk.[30]

Furthermore, whether harm is 'certain' to result from a given act is itself an artifact of how the question is framed-- in particular, the universe of acts over which we predict consequences and the epistemological POV from which we predict them. The failure to distinguish these framing decisions from the underlying causal relationship between act and consequences has led to considerable confusion in the literature. Three decisions are critical here: whether to take the small-number or large-number perspective; if we take the large-number perspective, whether to treat 'statistical certainty' as 'certainty'; and most importantly, at what moment to assess probabilities. I take up each of these issues in turn.

Small-number versus large-number perspective. Suppose we estimate that if we distribute a new flu vaccine, roughly one out of 0every one million people inoculated will have an adverse reaction, resulting in death. We plan to inoculate 10 million people. There are two ways to describe the likelihood that death will result from the inoculations, sometimes differentiated as "probability" versus "frequency." The first ("probability") is the odds that any given person who receives the vaccine will die (one in a million). The second ("frequency") is the total number of expected deaths if we inoculate 10 million people (ten). Both are

[30] Of course, some actions are overwhelmingly likely to result in harm each time they are done (viz, shooting a loaded gun held to someone's head). Most are not. But it is not clear what of philosophical interest follows from that observation. It just means that "virtually certain" is the limiting case on one end of the continuum of more or less risky conduct. Whatever criteria we use to distinguish acceptable from unacceptable levels of risk, it seems implausible that conduct on the two sides of that divide will be morally or factually discontinuous.

describing the same set of acts, predicted to cause the same consequences with the same likelihood; they are simply describing the likelihood from different perspectives.

The nonconsequentialist literature, however, frequently treats the two formulations as if they were describing different acts that entail different degrees of risk, and as a consequence carry different normative implications. Consider this recent treatment of the problem by James Lenman. Lenman argues that if the question on the table is, may an individual go out for a drive, knowing she is thereby putting others at some minuscule risk of harm, the answer is yes, in part because "the risks imposed are very small." But if the question is, what are the government's obligations in setting speed limits, we should apply a much stricter standard, in part "*because the small risks governments impose are typically imposed over very large populations,…so that it becomes an actuarial certainty some people will be harmed."*[31]

Read literally, the contrast Lenman sets up here is between prohibiting a given individual from driving entirely and permitting the government to regulate the manner in which all drivers drive. Surely Lenman will conclude that individuals are permitted to drive ***and*** that they are required to drive prudently, whether the question is posed from the perspective of what individuals owe to each other or from the perspective of how government should regulate driving.[32] The real

---

31 Lenman (2008, p.)(itals. added).

32 As I discuss elsewhere, some Kantians would distinguish these cases on a different basis: that self-regulation and regulation by the state are governed by entirely different norms. That position is (in my view) equally hard to defend, but for different reasons. See Barbara H. Fried, "The Consequences Matter—But to What?" (draft ms. 2010)

contrast he means to pose, I assume, concerns the level of care required in each case.[33]

To make this concrete, suppose the question is, how fast is it permissible to drive through a city street? Lenman's analysis suggests that he would specify a higher speed (say, 40 MPH) if the question is framed in terms of the duty of care *I* owe to others, and a lower speed (say, 25 MPH) if it is framed in terms of the level of care the government should legally require of all of us. The reason for the difference is that the odds are very low that *I in particular* will kill anyone if I drive at 40 MPH rather than 25 MPH (on a particular occasion? over a lifetime?), but it is a statistical certainty that, over a sufficiently large number of car trips governed by the speed limit, *at least one additional person* will be killed if the speed limit is set at 40 MPH rather than 25 MPH. Surely it cannot make sense to have anything turn on whether we consider identical conduct imposing identical risks in probabilistic or frequentist terms.

Is statistical certainty 'certain enough'? Virtually all acts, if measured over a large enough number of occurrences, are *statistically certain* to produce at least one death or serious injury. If 'statistical certainty' constitutes 'certainty' for

---

[33] The intuition that probabilistic and frequentist accounts of the likelihood of harm describe different underlying phenomena has given rise to a second confusion in literature: whether, when contractualists say that the "numbers should not count," they are referring to the frequentist presentation or the probabilistic presentation or both. Scanlon (1998), for example, expresses his opposition to letting the numbers count in frequentist terms ("The grounds for rejecting a principle are based simply on the burdens it involves, for those who experience them, without discounting them by the probability that there will be anyone who actually does so" (p. 208)), but then cashes it out with an example that relies on the probabilistic presentation (pp. 208-09). But just as it makes no sense to reach a different judgment about the wrongness of conduct depending on whether it is framed in frequentist or probabilistic terms, surely it makes no sense to let the numbers count if framed one way but not if framed the other. That implies that if there is *any chance* that someone will die from a given act, the ex post POV requires us to assess the complaint of that would-be potential victim at its full ex post value (that is, as certain death).

these purposes, then virtually all conduct is certain to result in someone's death, if we assess certainty from the large-number perspective. A few nonconsequentialists, including Lenman above, have taken that view.[34] Most have not, requiring some particularized knowledge about each instance of harm that will result (the identity of the victim, the precise nature of the injury) in order to treat harm as 'certain' to result. But for current purposes, the critical point is that either way we resolve the issue—statistical certainty is certainty or it isn't-- virtually all conduct will end up on the same side of the divide, pushing us either to a pure ex ante POV or a pure ex post POV across the board.

Knowledge is a dependent variable. Probabilities are epistemic facts. They describe what is known or knowable about future consequences as of the particular moment at which we assess them.

In the real world, the appropriate moment for assessment is determined by practical needs, and will generally be the present. That is to say, if we are trying to decide whether a "fail safe" valve on an oil rig is safe enough, we have compelling practical reasons to act on the basis of the best information available as of the moment we must choose in reliance on that information.

In contractualist thought experiments, in contrast, the appropriate moment to seek agreement on general principles to guide our actions is not determined by exogenous circumstances. It is internal to the contractualist argument and needs to be settled by appeal to normative arguments.

---

[34] For a particularly strong version, which not only treats statistical certainty as certainty but also treats statistical foresight as providing the mens rea necessary for murder, see Heinzerling (2006, pp. 521-34).

De facto, most contractualists have equated the moment of agreement on general principles with the moment at which a particular group of people must decide how to act in a particular situation pursuant to those principles—whether to flip the switch on the trolley, whether to rescue Jones, etc. When combined with the (itself unproblematic) Scanlonian commitment to let individuals act on the basis of whatever knowledge they are imagined *actually* to possess at the time of agreement (what Sophia Reibetanz has termed the 'natural veil of ignorance'), the result is a POV that is contingent on the happenstance of the factual posture in which a real or hypothetical dilemma arises. If, as in Kamm's Ambulance I, each person must decide whether to allow the ambulance to speed at time when she has no reason to think that *her* expected outcome deviates from anyone else's in the town, it would be reasonable for her to decide to go by the numbers. If she has to revisit the issue in Ambulance II, when it 'just so happens' that the ambulance is about to mow her down, we throw out her prior agreement as no longer reasonable to hold her to, and let her choose on the basis of her updated information about her self-interest. That is the burden of Lenman's argument that a policy that "is not acceptable at every time [to every person] is plausibly acceptable at none."

It is hard to overstate the argumentative consequences of contractualists' having de facto equated the moment people must agree on general principles with the moment that some subset of them must choose how to act in a particular situation pursuant to those principles (the "here and now," in Kamm's words). To put it in the broadest terms, if contractualists were to decide on further thought that the only morally appropriate moment to ask for agreement on general principles

governing the risks we may impose on each other is *before* any of us knows exactly what risks we will face and how they will play out for us-- to conclude, that is, that agreement must be reached from some version of a truly ex ante POV-- then by contractualists' own acknowledgement, regulation of virtually all socially useful but potentially harmful conduct may be (must be?) turned over to what amounts to some form of interpersonal aggregation, in which the numbers—both the numbers of potentially affected individuals and the probability of harm or benefit to each of them—count. A few questions will remain in the nonaggregationists' domain: Are there some future interests that individuals should not be allowed to put at risk? What do we do about individuals who are in a position to know *ex ante* that the expected outcomes for them are worse than for others?[35] But all the remaining moral quandaries are swept away. No more trolley problems, no more Lifeboat problems, no more Ambulance Cases, no more technicians trapped under transmission wires, because we are no longer interested in what general principles someone would agree to once she knows for a fact that her own life is at stake in a particular choice that will be governed by those principles. The only morally relevant question is, what would she have said *before* she knew that?

If, on the other hand, contractualists really mean to insist on the moral imperative of an ex post POV—insist, that is, that before deciding whether to adopt a

---

[35] For discussion of the latter issue (the so-called Amish problem), see section 2.2 infra. A number of other factors going to the level of precaution taken by the actor and the victim survive the shift to an ex ante POV. But, as I discuss at length below, it is not clear that the requirement of due care is cashed out in a way that is distinct from the optimal levels of precaution that would be mandated under straight aggregation.

given principle, *everyone* has a right to know how things will turn out *for them in their actual lives* if it is adopted-- the result (inescapably) would be moral gridlock.

But instead, contractualists have, de facto, split the difference. They have equated the moment people should agree on a general principle with the moment some subset must resolve an exigent crisis pursuant to that principle, and then endowed that subset with whatever knowledge each of them *happens to possess at that moment* about how things will turn out for them if the exigent crisis is resolved in accordance with the proposed principle. The POV that results is an odd hybrid of ex ante and ex post. If, at that moment, someone happens to know that adopting Principle X will result in her certain death, her complaint about principle X will carry the full disvalue of death (Kamm's Ambulance II). If she knows only that she, like everyone else, faces a 1 in 1 million chance of death, her complaint will be discounted by that probability (Kamm's Ambulance I).

If, at that moment, everyone happens to be 'certain' about their fate (the typical factual posture of hypotheticals in the philosophical literature on harm to others), under the minimax complaint rule adopted by Scanlon and others, we end up with whatever principle will produce the least-bad worst outcome. If everyone faces the same certain worst outcome (e.g., death), some other supplementary rule is needed to break the tie (e.g., Scanlon's tiebreaker rule, which would let the numbers count in that situation, the act/omission distinction, which gives the victory to the pedestrian in Kamm's Ambulance II).

If, on the other hand, everyone happens to be uncertain about their fate and everyone faces identical risks, we would expect to end up with a principle that

approximates the optimific one—that is, the one that yields the best aggregate outcome (the 'right' result in Ambulance I).

But if the parties have asymmetric information about their fate, the hybrid rule produces outcomes that even its proponents recognize to be unacceptable. The problem has been much discussed in the context of Sophia Reibetanz's "Unexploded Mines" hypothetical:

> One hundred workers are working in a field, in which an unexploded mine is known to be located. A nearby person, X, is the only one who can disarm the mine. If X disarms it, X is certain to get pneumonia; if X doesn't disarm it, 1 of the 100 workers "is certain" to be seriously injured, but it is unknown which one. The injury to that one worker will be 10 times worse than pneumonia will be for X.

Under the hybrid rule, Scanlon's Relative Complaint Model produces the unacceptable answer that X doesn't have to disarm the mine.[36] Because X knows for certain what will happen to him ex post if he disarms the mine, X's complaint about pneumonia gets weighted at the full disvalue of pneumonia. On the other side, since "it is not known, in advance, which of the hundred laborers will encounter the mine,"

---

[36] Reibetanz (1998). A cognate problem arises in accounting for the interests of future generations who will be affected by whatever general principle we adopt. As Parfit has noted, if we take seriously Scanlon's 'individualist' restriction, the future complaints of those not yet existent persons shouldn't account at all. (2011, p. [ ]). But even if we posit representative future persons to stand in for the interests of actual future persons, the complaints of each of those future victims would have to be discounted to reflect the very low probability that any given one of them in particular will suffer from our present actions, Given that the interests of presently existing persons in choice of a given principle will generally be more certain than those of future generations, the hybrid rule will systematically overweight the interests of present generations relative to future ones. This is clearly not contractualists' intent. But it is not clear how one can avoid it under the hybrid rule that requires us to weight complaints of others only insofar as we actually know them to be at risk of being harmed by us.

each laborer's complaint about the potential severe injury to him is discounted by odds he will incur it, yielding a complaint of 1/100 x 10 x, which is $1/10^{th}$ of X's.

Contractualists have suggested a number of ad hoc fixes to solve the problem. But the problem itself is, I believe, a non-problem, resulting from the peculiar way that contractualists have, de facto, constructed the choice situation from which a hypothetical agreement on general principles is sought: Locate the moment of agreement on a general principle to govern X choice at the moment some arbitrarily selected group of people must make an exigent X-type choice in accordance with it; limit the representative persons whose complaints are to be weighed in the balance to the members of that group; and then weight each of their complaints about candidate principles based on whatever information they just happen to possess *at that moment* about how things will turn out for them if a given principle is adopted.

What is the right choice situation from which to seek hypothetical agreement on general principles? I don't think there is a simple, or single, answer to the question. . It depends in part on one's normative commitments, in part on the nature and scope of the principles in question. But the answer that much of the contractualist literature has, de facto, adopted seems to me singularly hard to defend. At a minimum, it needs to be defended.

2.1.2. Right Result, Wrong Reason.

None of this is to deny that most people respond very differently to acts that impose virtually 'certain' harms on others (or rescue them from virtually certain harms) than they do to acts that expose others to a 'mere' risk of harm. If that

difference can't be explained by moral or factual distinctions between 'certain harm' and 'risk,' what does explain it?

In What We Owe to Each Other, Scanlon offers one answer that he has since partially repudiated. In brief, Scanlon argues that acting in a fashion that is "sure" to result in harm is close to conclusive evidence that one acted without due care for the interests of others, and hence acted impermissibly. In contrast, acting in a fashion that imposes a mere risk of harm is permissible, provided one takes "reasonable precautions" against harming others. Any more demanding standard, Scanlon concludes, would be "too confining." Thus, Scanlon arrives at roughly the same hybrid solution but for very different reasons. Scanlon's intuition here seems to be that if harm is absolutely foreseeable, it is avoidable, and the failure to avoid it is per se wrongful. But of course uncertain harms can be avoided as well, by avoiding the conduct that imposes the (just as foreseeable if lower) risk of harm. The question in both cases is whether we want to pay the price of avoidance—a question that has to be answered on other grounds.

The intuition that imposing 'certain' harms is different in kind from imposing a mere 'risk of harm' is better explained, I believe, by the emotional and psychological valence of immediacy and certainty. Contemplating imminent death to an identified victim tends to concentrate the mind wonderfully on the horror of causing or permitting it to happen. For the victim facing imminent death, the horror of knowing that others could save her but choose not to is likely swamped by the prospect of death itself. But the victim is not the only (or indeed maybe even chief) party whose *sensibilities* are at risk in such cases. There is the actor herself,

who, in order to follow through on a prior agreement, must now knowingly kill (or fail to rescue) an identified victim. There are also the bystanders who must witness an "intentional" killing or avoidable death up close and personal. And when that death or serious harm can be prevented with mere money or the trivial loss of pleasure -- Baby Jessica in the well, Scanlon's Jones trapped under transmission wires-- it seems more than a little cold, maybe downright inhuman, to refuse to incur those costs simply because, in the aggregate, the costs are large.

This completely natural response, I think, explains Scanlon's emotionally charged description of the resulting death in a case like Ambulance II as an "intentional killing." But in what sense intended? Clearly it is not intended in the strong sense of 'desired'; it is intended in the weak sense of 'foreseen'. But the same is true of any accidental deaths that might result from the agreement in Ambulance I, absent any opportunity to rescind one's consent. What differentiates the cases is not intent per se but the greater specificity and proximity (and, from some perspectives, updated probability) of the foreseeable harm.[37] Scanlon, believing those factual differences make a dispositive moral difference, describes the resulting deaths in Ambulance I as merely accidental harms and those in

[37] If the hypothetical agreement in Ambulance I encompasses only the residents of one town, the likelihood that *any* pedestrian will die will over the next x years as a consequence of the agreement may be very low, given the low number of trips the ambulance will take. In that case, the very likely impending death in Ambulance II would force us to revise upward the probability of *someone* dying. But that result is the artifact of our choice to limit the scope of the hypothetical agreement to this one town. If the agreement on principles in Ambulance I and Ambulance II covered (say) the entire United States, over some period of time the statistical certainty of at least one death would approach 1.0.

Ambulance II as intentional killings. Others, concluding they make no moral difference at all, describe both, equally, as 'knowing killing' or 'murder.'[38]

Perhaps, as some have argued, our heightened emotional response to imminent, identifiable harms itself has a moral basis.[39] Perhaps it is explicable only in psychological or emotional terms. In a practical sense, it may not matter. If, for whatever reason, it is the case that *none* of us will want to stand by and do nothing while the driver in Ambulance II elects to mow down one pedestrian to save five passengers pursuant to the agreement reached in Ambulance I (or worse yet, to find ourselves to *be* the driver who has precommitted to mow down the pedestrian) and those feelings will be strong enough, in the event, that we will refuse to abide by a prior agreement to the contrary and condemn anyone who doesn't similarly

---

[38] See Heinzerling (2006) (arguing that statistical certainty is sufficient to impute intent in the strong sense of the mens rea required to characterize the resulting deaths as 'murder'). Cf. Lenman (2008).

[39] Kant put a moral spin on the overwhelming impulse to renege on a universal law of nonbeneficence when one's own survival is on the line: "'[A] will which decided in this way would be in conflict with itself, since many a situation might arise *in which the man needed love and sympathy from others*, and in which, by such a law of nature sprung from his own will, he would rob himself of all hope of the help he wants for himself.'" (Example 4, Groundwork, quoted in Herman (1993, p. 47). More recently, Sarah Miller has argued that our greater concern for identified and proximate victims reflects an ethics of care, grounded in the face-to-face interpersonal relationship between the rescuer and the rescued. Meeting those needs is "inherently interactive" between the "carer and cared," and the "particularity" of the person in need is essential to engage our "emotional capacities of empathy and sympathy, moral skills that aid in establishing identification with the plight of others." Miller (2006, pp 137-160, 141-143). While Miller's argument is addressed only to the Duty of Rescue, a fortiori the same considerations should lead her to support a heightened duty to avoid harming others when the would-be victims are identified and proximate. See also Lenman (2008, p. 116); Parfit (2011, p. [ ]): "When we know that the lives of certain people are in danger, as would be true, for example, if some group of miners are trapped underground, we have reasons to want great efforts to be made to save these people's lives. Some economists point out that we would do more to increase people's life-expectancy if, rather than spending huge sums on trying to save known particular people in such emergencies, we spent this money on more cost-effective safety measures that would prevent a greater number of statistically predictable future deaths. But we could reasonably deny that this fact is morally decisive. We have strong reasons to want great efforts to be made to save the lives of known particular people who are in danger. By making or supporting such efforts, for example, we reaffirm and express our solidarity with, and concern for, everyone in our community. That is less true of acts that merely prevent the statistically predictable future deaths of unknown people."

repudiate it, then such feelings have to be taken into account in some fashion, if we are to cobble together a stable policy response to the problem. But how exactly we should take them into account depends (inter alia) on whether we think we are dealing with two independently valid responses to different questions or simply temporally inconsistent responses to the same one.

2.1.3. A Pyrrhic Victory?

Resolving uncertain harms from an ex ante POV and certain ones from an ex post solves the moral gridlock problem for most real-world problems, for the simple reason that in the real world, people generally must commit to a course of action while the outcome is still uncertain on all sides, and rarely have a chance to rescind their consent once their own fate becomes clear. It is easy to miss this truth, coming at the problem through the lens of the contemporary philosophical literature on harm-to-others, because of the prominent role given to hypotheticals involving 'certain' harms (trolley problems, etc.) and the marginalization of the problem of uncertainty (risk).

One of the unfortunate consequences of this inverted picture of reality that emerges from the philosophical literature, I believe, has been to encourage contractualists to think that if they cede the problem of risk to aggregation but require "certain harms" to be resolved by nonaggregative principles, they will have carved out a meaningful middle way—one that retains a substantial role for nonaggregation while avoiding moral gridlock. In fact, under the 'we get certainty and aggregationists get uncertainty' compromise, almost all the real work of regulating potentially harmful conduct is ceded to aggregation. The odd assortment

of one-off rescue and 'certain harm' cases that populate the philosophical literature—trolley problems, Jim and the Indians, Ambulance II— in which we must choose between two 'certain' bad outcomes are, in effect, just a luxury concession to our nonaggregationist intuitions.

I do not think it could be otherwise, because the implications of nonaggregation in the context of harm to others will be palatable only as long as the occasions for deploying it are rare. Whether or not one agrees that it is morally wrong to kill one identified person to save five, the fact is, social life will go on pretty much as before if every time we encountered that dilemma in its various postures, we let the five (or five hundred or five thousand) die. But if 'certain' harms were as common as accidental ones, the social costs of refusing to let the numbers count would be morally unacceptable to all.

Scanlon's response when the going gets tough under his Relative Complaint Model leaves little doubt where he would come out. Scanlon recognizes that if the maximin rule built into that model were applied to all activities that imposed a risk of serious harm on someone, we could not (just for starters) build sports stadiums or skyscrapers, drive cars, fly airplanes, because the risk of death to bystanders (should it materialize) would clearly trump the inconvenience or foregone pleasure that would result from banning such activities.[40] But, faced with that implication, instead of biting the bullet Scanlon abandons his Relative Complaint Model as "too

---

[40] There is another out available to Scanlon, under his "Tie-Breaker" principle. That principle provides that where the worst possible outcomes under two competing alternatives are identical, the numbers can count as a tie-breaker. Scanlon (1998). As I discuss below, one could plausibly argue that the worst possible consequence of all conduct (albeit a very remote one in many cases) is death, turning all harm-harm comparisons into ties. But this renders the Relative Complaint Model superfluous; all risky conduct is turned over to aggregation, in two steps rather than one.

confining."[41] In its place, Scanlon proposes that we reach a reasonable compromise among competing interests, one that takes into account both our interest in "avoid[ing] bodily injury" and also the "cost[s] from the point of view of potential agents" of "a general prohibition on … acting" in a fashion that would put others in harm's way.[42] In short, we should accept the central principle that the "individual reasons" restriction was meant to rule out: that relatively trivial ex post benefits to one group can offset serious costs to another, provided the former group is large enough and the latter group small enough.

2.2. Allow Disadvantaged Parties to Reject an Unequal Distribution of Ex Ante Risks.

As I noted above, individual agreement sought from an ex ante POV will tend to approximate the aggregative solution, provided that everyone faces roughly the same *expected* costs and benefits from a proposed activity, given their individual preferences. What if they don't? Recent discussion of this issue has focused on the following hypothetical.

A group of Amish farmers live under a flight path, putting them at heightened risk of falling planes. The flight path cannot be relocated, except at a prohibitive cost. The Amish don't face any greater objective risk than others also living under

---

[41] Scanlon (1998, p. 209). For a trenchant discussion of Scanlon's inconstancy to the Relative Complaint model in the context of the duty to rescue, see Ashford (2003, pp. 298-230). It is not clear to me whether Scanlon and other contractualists mean the Relative Complaint Model to apply in the same (strict) fashion to the affirmative duty to aid that it ostensibly does to the negative duty not to harm. If they don't, then the inconstancy Ashford notes need not constitute an internal contradiction in Scanlon's theory. But it seems hard to avoid the conclusion that it is internally contradictory in the context of harm to others.

[42] Scanlon (1998, pp. 204, 205).

the flight path, but unlike the others, they derive little or no benefits from air travel, either directly or indirectly.[43] Thus, the *expected value* to them of permitting planes to fly is much lower than everyone else's, and may well be negative (although trivially so, given the minuscule probability that a plane would fall on them). Do they have reasonable grounds to object to the current flight path?

The distributional problem posed by the Amish hypothetical arises frequently in the real world in less fanciful forms. Land use regulation is a common site of such conflicts, because the negative consequences of a particular land use are often geographically concentrated while the benefits are more widely diffused. A standard case here would be the siting of a toxic waste dump. Placing the dump near a poor community (Poorville) may well be optimal from an aggregative (welfarist) perspective, because the dump will typically depress land values in Poorville less than in nearby richer communities. But because they will bear a disproportionate share of the risks from the toxic dump and at best only a pro rata share of the benefits from whatever activities generated the waste, the residents of Poorville could reasonably object that the aggregatively optimal solution is unfair *to them*.

A Scanlonian contractualist is committed to taking those objections seriously. But what exactly does that entail? We could (and sometimes do) require those disproportionately benefitted by siting the dump in Poorville (the company

---

[43] See Munoz-Darde (2009, p. [ ]); Lenman (2008, p. 121, n. 40); Ashford (2003, pp. 298-99). The example is a variant of a hypothetical originally raised by Scanlon, in which the group facing a disproportionate share of the downside risks was the poor rather than the Amish. Scanlon (1998, pp. 208-209). For Scanlon's own skepticism that such cases arise frequently, see id.

generating the waste, the larger polity that gets at least their fair share of the benefits without bearing any of the costs) to compensate Poorville for expected or actual harm. But from a contractualist perspective, this only partially addresses the complaint of Poorville, which (by hypothesis) does not want the toxic dump there at all. Where the potential harm to Poorville is severe and not fully compensable by money, Poorville might well refuse to agree to the siting for any amount of money (or at least any amount anyone is likely to offer them).

Now what? Broadly speaking, there are three possible responses. The first is to force Poorville to accept the toxic dump along with cash compensation. Writ large, this amounts to a welfarist regime with strict Paretian constraints. That is to say, we decide what potentially harmful activities people may engage in by some aggregative procedure, but require the losers to be compensated for their loss. My guess is that most contractualists will resist this solution, at least for serious risks to health (at the extreme, death). If not, then they have ceded the problem of risk to aggregation full stop, leaving only the question of compensation to be resolved by nonaggregative principles.

Second, we can judge the fairness of the ex ante distribution of risks over some larger class of activities. If, over some suitably broad class, the (expected) losers on one risky project are likely to be the (expected) winners on another, we can treat those two events as offsetting. I take up this alternative in 2.3 below. For now, I note only that where death is one of the risks, many contractualists will balk at that solution.

Third, we can insist that Poorville be given a disproportionate say in the initial siting decision, at the extreme a veto over it. This solution seems more in the spirit of Scanlonian contractualism. It may well be both appropriate and doable in discrete cases, where the social costs of giving the losers a veto are manageable and the expected costs to them if the project goes forward are severe (health hazards, destruction of a community, etc.). Again, land use disputes are a prime example here. That the losers in such cases are often the poor strengthens the argument, in my view and the view of many others, for giving them a veto on fairness grounds.

But contractualists' willingness to bite the nonaggregationist bullet in these cases is likely to run out pretty quickly. Many large-scale construction or public works projects concentrate expected costs on one group, and many of the expected costs are not fully compensable with money. If every disproportionately burdened group were given a veto over the project in question, we would be back to gridlock pretty quickly. Giving the group a disproportionate say over the principle but something short of a veto would lessen the risk of moral gridlock, but at the cost of leaving the group still vulnerable to being sold out to benefit the majority. And of course giving any voice less than a veto raises the question of degree: should the minority get two votes for every one of the majority? Ten votes? How should we go about deciding this? It is notable that Scanlon's willingness to give the expected losers a veto in such situations runs out almost immediately. Having posed a variant of the Amish problem himself, Scanlon concludes that of course air travel cannot be banned, because to do so would be "too confining," notwithstanding that

his own Relative Complaint Model of reasonable rejection pretty unambiguously requires that result.[44]

Finally, it is critical to keep in mind the limited reach of this exception to aggregation, even in theory. It is triggered only when the *expected* net benefits of an activity are distributed unequally. It does not apply where expected costs and expected benefits are distributed roughly equally, but ex post there will big losers. Most accidental harms fall in the latter camp.

2.3. Aggregate Risks *Intra*personally.

Because the Amish pose such an extreme case, their complaint that they are bearing the risks of social progress without expecting to reap any of the rewards might well hold, however we bundle risky activities that might potentially affect them over a lifetime. But in the typical case, whether the expected net benefits of risky conduct are unfairly distributed may depend on whether we evaluate each act (or type of act) authorized by a given principle separately or bundle them in some fashion.

Arguments for bundling (henceforth, ***intra***personal aggregation) have been around for a long time with respect to compensation for harm.[45] While the technique has gotten less attention in setting acceptable levels of risk, those who

---

[44] Scanlon (1998, p. 209).

[45] The seminal article here is Fletcher (1972). The literature focusing on *compensation* rather than prohibition is frequently ambiguous on whether it is adopting an ex ante or ex post POV, starting with Fletcher's article. Most of Fletcher's argument takes the ex ante perspective, focusing on the maldistribution of risk rather than of harm. But some comments suggest this is premised on the assumption that reciprocal risks will produce equal harms over an appropriately long time horizon—that is, that ex ante equals ex post. See id. at 547: "[I]t would make little sense to extend strict liability to cases of reciprocal risk-taking, unless one reasoned that in the short run some individuals might suffer more than others and that these losses should be shifted to other members of the community."

have considered it generally support it, and in my view it is difficult to defend *not* bundling in some fashion.[46]

Aggregating risks ***intra***personally would undoubtedly change contractualists' judgment about the reasonableness of risky conduct in some number of cases. How large that number is depends on how the aggregation is to be done, an issue not addressed to date in the contractualist literature. (Over what classes of activities? Over what period of time? Do we include all expected outcomes or only those that exceed a certain threshold amount?) But however large it is, it won't solve the basic problem facing nonaggregationists: that an ex ante POV proves too little and an ex post POV too much. It simply replicates it across bundled groups of activities.[47]

At the loser tail of the distribution, we still have the Amish problem: What do we do about individuals who face huge expected losses relative to everyone else?

---

[46] Lenman (2008, p. 108); Reibetanz (1998, pp. at 299-300) ("since it seems that the morally relevant unit is not the isolated instance of adherence to a principle, but the lifetime of acceptance of it, we should add that the effects relevant to an individual's complaint about a principle are the effects of general acceptance of that principle over her lifetime."); Scanlon (1998, p. 237). While expressing skepticism that this argument will take care of conflicting natural duties, Rawls seems to embrace it with respect to the basic structure: "Since we are dealing with a comprehensive scheme of general rules, we can rely on certain procedures of aggregation to cancel out the significance of the complicating elements of particular situations." Rawls (1971, p. 299).

Here, as elsewhere in the literature, authors are frequently ambiguous as to whether the problem they are trying to solve is ex ante risk imposition or ex post compensation. See, e.g., Hansson (2007, pp. 31-32).

[47] Lenman suggests that we limit use of intrapersonal aggregation to "mildly risky behavior," which he defines as behavior that "impos[es] small risks on others." Lenman (2008, p. 108). This solution appears to straddle ex ante and ex post POVs. Assessed from a true ex ante POV, most socially beneficial activity imposes at most a small risk on any given individual. If instead Lenman means to measure the severity of risk by the severity of the worst expected outcome, then he has effectively jumped to an ex post POV.

The only difference is that now we are measuring each person's expected losses over a lifetime rather than activity by activity. Giving expected lifetime losers a veto over conduct that, cumulatively, puts them at a disproportionate risk of harm takes on board all of the difficulties of giving such a veto on an activity-by-activity basis. Are we really prepared to prohibit the conduct in question rather than compensate the losers? It faces some additional problems as well. As we proliferate the risks that have to be weighed, the informational burdens grow exponentially. Except in outlier cases like the Amish, we are very unlikely to be able to identify ex ante the *expected* lifetime losers from the potentially harmful conduct of others. And, if the losers are to be given a putative veto, how do we decide which of the many risks that cumulatively leave them expected lifetime losers to prohibit on their behalf?

Finally, as in the activity-by-activity assessment of fairness, this approach provides a viable alternative to interpersonal aggregation only when *expected* lifetime harms are unfairly distributed. It does not address the more typical case in which everyone faces roughly comparable lifetime risks of bodily harm ex ante, but ex post only a few will be big lifetime losers.

2.4. Take Into Account All the Consequences of a Principle.

What contractualists are seeking agreement over is not rules governing individual actions but rather "general principles of action."[48] In Scanlon's words, in assessing such principles this requires us to take into account "not only the consequences of particular actions, but also the consequences of general

[48] Scanlon (1998, p. 171); Wood (2011, pp. [ ]): "One thing I hope is clear by now is that for Kantian ethics, the point of a moral principle… is not directly to tell us what we should *do*. It is rather to ground a set of rules or duties, and more generally to orient us as to how we should and should not *think* about what we should do."

performance or nonperformance of such actions." By this, Scanlon intends not just that we should aggregate each individuals' complaints *intra*personally over the individual's lifetime or some other significant period of time, but also that we take into account the "cost[s] from the point of view of potential agents" and society at large of "a general prohibition on … acting" in a fashion that would put others in harm's way.[49]

One of Scanlon's motivations for insisting that what we are after is "general principles of action" rather than rules governing particular actions is precisely to avoid the moral gridlock problem.[50] But whether the compromise Scanlon suggests-- take into account both the potential costs to those who will be harmed by risky conduct *and* the potential benefits to the actors and society at large of engaging in such conduct-- supplies a genuine middle way between gridlock and aggregation or instead is a covert capitulation to aggregation full stop depends on how we are to take the costs and benefits into account. Scanlon doesn't say enough here to reach a judgment one way or the other. But what he *does* say sounds sufficiently like conventional interpersonal summing of costs and benefits that, I think, it is fair to put the burden on Scanlon to explain how it differs.

Or consider this example from Parfit. Parfit argues that Scanlonian contractualism can save us from the answer utilitarians must give to the standard 'Forcible Transplant' hypothetical: May (must??) a surgeon kill one patient if by doing so he can harvest the patient's organs to save five? Utilitarians, on these

---

[49] Scanlon (1998, pp. 203-205, 264, 273).
[50] Scanlon (1998, p. 170).

facts, must say yes, says Parfit, because forcible transplants will produce the optimal outcome. Scanlonian contractualism saves us from that obviously wrong conclusion, argues Parfit, by allowing us to count on the other side of the ledger the "anxiety and mistrust" such a regime would induce in all of us: "[T]he saving of these other lives would be outweighed by these ways in which it would be bad for us and others if, as we all knew, our doctors believed that it could be right to kill us secretly in this way."[51] *But this is exactly the argument that any sensible (act or rule) utilitarian would offer to explain why killing one to save five in this context is inadvisable from a utilitarian perspective.*[52] Is Parfit weighing the short-term benefits (in lives saved) and long-term demoralization costs differently from how a utilitarian would weigh them? If so, what is the difference?

2.5. Tiebreakers Go to Ex Ante POV.

Since Taurek's famous 1977 article, nonconsequentialists have been wrestling with ways to justify letting the numbers count, at least as a tiebreaker when we are forced to choose between qualitatively commensurate harms.[53]

The tiebreaker strategy has figured prominently in recent contractualist

---

[51] Parfit (2011, p. [ ]).

[52] Parfit appears to believe otherwise, at least for act utilitarians, because they would (be required to?) "consider Transplant on its own" and ignore the systemic costs if such deliberate killings became widespread and hence widely known. (2011, p. [ ]). This seems to me wrong. A pure act utilitarian might, I suppose, support one Transplant if she were sure the doctor's actions would remain secret. But Parfit is supposing that if repeated frequently enough, Transplant would become public, with all the resulting demoralization costs. Once Transplant is public, act utilitarians would cease to support it, and until it is public, Parfit (at least here) has supplied no reason why contractualists should not support it. To put it another way, if you ask act utilitarians whether they would support one act of Transplant, they might say yes, on the 'one free bite at the apple' theory. But if you ask them whether they will support 1000 such acts, they would say no for precisely the same reason that Parfit says no.

[53] Taurek (1977). For just a few of the many responses to Taurek's dilemma, see Otsuka (2000); Kamm (1993, chs. 5 and 6); Parfit (1978); Kavka (1979); Sanders (1988); Scheffler (1994, p. 119).

accounts.[54] Whether it offers a viable middle way is unclear. First, there is the question of scope. Almost all conduct carries with it some risk, however remote, of death to someone; outside of the world of philosophical hypotheticals, no conduct carries with it a non-statistical certainty of death. If (as we are required to do under a true ex post POV) we ignore probabilities, every case is a tie, because every case could result, in the worst-case scenario, in death. At that point, we have aggregation full stop, in two steps rather than one. We can avoid that conclusion by letting in probabilities in some fashion, but if we do that we no longer need a tiebreaker rule to let the numbers count. We have done it directly, thereby giving up on contractualists' central commitment to 'individual reasons.'

Assuming one could come up with a principled way to keep some but not all cases outside the jurisdiction of the tiebreaker rule, any cases that are outside revert to the ex post POV, in which the numbers do not count in any way. In theory, this could provide contractualists with a viable middle way. The question is whether they can live with its implications. The tiebreaker solution limits the unpalatable consequences of an ex post POV by limiting the cases to which it applies. But unless the cases that fall outside the tiebreaker solution are few in number and trivial in social importance, I am doubtful that contractualists will be willing to refuse to let the numbers count even there. And if the cases that fall outside the tiebreaker rule *are* trivial compared to the cases that fall within it, it raises the question whether the tiebreaker rule should be considered a meaningful middle

---

[54] Scanlon (1998); Kamm (2007, p. 34); Reibetanz (1998); Parfit (2011).

way at all. Like the distinction between 'certain' and 'uncertain' harms, it begins to look like just a luxury concession to our nonaggregationist intuitions.

2.6. Avoidability of Risk Waives Individual Complaints.

Another path to a middle way that contractualists have explored is to distinguish between permissible and impermissible conduct based on whether the victim could have avoided the risk.[55] In What We Owe, Scanlon helpfully distinguished between two versions of the argument: the Forfeiture View and the Value of Choice View.

The Forfeiture View corresponds to the more familiar legal doctrine of "assumption of risk." The basic intuition here is straightforward. If a party knows about the risk he or she faces in pursuing an activity and pursues it nonetheless, she tacitly consents to bear any untoward consequences.[56]

The problems with attaching so much moral significance to tacit—or indeed explicit—consent have been extensively discussed in the choice-egalitarian literature. The most serious one, as choice egalitarians themselves acknowledge, is that assuming a risk knowingly and voluntarily is not sufficient to establish one's moral responsibility for the outcome, whatever it might be. Most risks can be avoided--—if nothing else, by avoiding the activity that imposes them. We

[55] It is not always clear whether this argument is addressed to the permissibility of the risky conduct or the obligation to compensate victims. My comments should be read to apply only to the former.
[56] For one defense of the moral relevance of tacit consent, see Ripstein (2009, p. 47).

therefore have to ask whether it would be reasonable to ask the victim to take whatever action is necessary to avoid the risk.[57]

Scanlon's Value of Choice model responds to this last concern, among others, by shifting focus from the 'voluntariness' of the choice to the "value of the opportunity to choose that the person is presented with." In Scanlon's words, "If a person has been placed in a sufficiently good position, this can make it the case that he or she has no valid complaint about what results, whether or not it is produced by his or her active choice."[58] Focusing on the state as actor, Scanlon argues that a person will be placed in a sufficiently good position if "officials have 'done enough' to… reduce the likelihood that anyone will be injured."[59] One of the things officials can do is to inform would-be victims of the danger so they can take reasonable precautions to avoid the risk. Another is to take reasonable precautions themselves to minimize the risk.

Whether Scanlon's Value of Choice model offers a solution that is distinct from aggregation once again depends on exactly what it means to "do enough" or take "reasonable" precautions. Welfarists would answer the question by finding the optimal tradeoff between the costs and benefits of safety precautions. As discussed in more detail in 2.7 below, it is unclear whether contractualists would answer the question differently, and if so, how.

---

[57] For recognition of this problem in the context of compensation, see Fletcher (1985, p. 84).
[58] Scanlon (1998, p. 258).
[59] Scanlon (1998, p. 258).

2.7. Require Greater Safety Precautions Than Under Straight Aggregation.

In the typical case of risky but socially useful conduct, the choice we face is not whether to permit risky act X or forbid it, but rather what level of precaution to require as a precondition of doing X. (To put it another way, forbearing from doing X at all is the limiting case of the available levels of precaution.) A number of contractualists have suggested that the middle way lies in requiring a higher level of precaution when undertaking potentially harmful conduct than would be required by a straight summing of costs and benefits.

In What We Owe to Each Other, Scanlon suggests one way we might do that: Rather than permitting less precaution the greater the social benefits of the activity (as he takes aggregation to do), we should demand the same level of precaution irrespective of the benefits to society. [60] The contrast Scanlon sets up here is based on a misunderstanding of how aggregationists would think about the problem. In standard aggregative techniques, three numbers are relevant to determining the optimal level of safety precautions to employ if activity X is to be done: the social benefits of X; the expected harms if X is done with a given level of precaution; and the costs of undertaking that level of precaution.[61] If, however much we spend on safety, the sum of lives lost plus safety costs will be greater than the social benefits, it doesn't make sense to go ahead with the project at all. In answering that

---

[60] "We do not think that a higher level of safety must be provided for workers on a building that will benefit only one family as opposed to an apartment house or a public bridge." Scanlon (1998, p. 236).

[61] Lives lost can be expressed as the probability of any given person dying or the total number of expected deaths. Contractualists have often treated these as describing a different set of material consequences, but (as discussed above) they are simply a different way of describing the same material consequences.

threshold question only—whether it makes sense to go ahead at all-- , the benefits of the project *are* relevant to the level of safety precautions, but they are directly, not inversely, correlated: All else being equal, the greater the benefit of an activity the greater the costs (including money spent on safety precautions) we will tolerate before pulling the plug on the project entirely as a net social loss. But once it is established that there is some level of safety precautions at which it will increase social welfare to go forward with X, the benefits of the project are irrelevant to calculating the optimal level of safety precautions. For a straight aggregationist, the optimal level depends solely on the marginal tradeoff between money spent on safety precautions and increased safety. The optimal level of safety precautions is reached when the incremental cost of additional precautions will exceed the benefit (in increased safety) they will produce.

Once contractualists conclude (on whatever basis) that at some level of safety precautions it is reasonable to permit a particular risky activity to go forward, they are in exactly the same position as aggregationists.[62] The level of expected benefits is no longer relevant. The only question on the table is how to trade off the costs and benefits of different levels of precaution. If contractualists can generate an alternative to aggregation by varying the required level of precaution, it has to come from their making that tradeoff differently.

---

[62] It is not clear that Scanlon would make that threshold determination any differently from aggregationists. "In each case, in order to defend the practice in question we need to argue [first] that the importance of the social goal justifies creating the risk . . ." (1998, p. 264).

By way of illustration, consider the following. A city is deciding whether or not to put up a new sports stadium. The city estimates that the stadium will generate subjective entertainment value for spectators equal to $100/person per year. An expected 100, 000 persons will attend games per year, for a total expected benefit of $10 million per year.

If $1 million is invested in safety precautions during construction to protect passersby, four pedestrians are expected to die. If the amount is raised to $10 million, expected deaths are reduced to two. If it is raised $50 billion, expected deaths are reduced to one. No amount of money spent on safety precautions will reduce expected deaths to zero.

If we evaluate this project from a true ex post POV, a Scanlonian would have to conclude the city may not build the stadium, whatever level of safety precautions it takes, because (peeking ahead) the complaint of the one passerby who will be killed will outweigh whatever amount of pleasure spectators will derive in the aggregate over the life of the stadium. As discussed in section (1) above, however, Scanlonians reject a pure ex post POV if the potential harms to others are uncertain at the time the decision must be made whether to build the stadium. Instead, if the social benefits of putting up the stadium are great enough, they will permit it to be built notwithstanding the non-zero risk of death to passersby, provided reasonable precautions are taken.

What level of precaution is "reasonable"? There is a tendency among contractualists and other nonconsequentialists to think that if we could do anything

more to reduce expected deaths, it would be "unreasonable" for us not to do it.[63] In the stadium project, that would mean it is unreasonable for the builder to spend any less than $50 billion on safety precautions. But in most arenas of life, as undoubtedly in this project, the point at which further investments in safety will cease to produce any positive return in safety is far beyond the projected benefits of the project. The result of interpreting "reasonable" in this fashion will thus be moral gridlock by a different route. No building project will show a net expected benefit, given the enormous amount one would be required to spend on safety precautions, and hence no building project may go forward.

So, presumably, there is a point at which even contractualists will conclude that an additional investment in safety will produce so little return in increased safety that it is unreasonable to require that additional investment, and the project may go forward without it. As noted above, for aggregationists that point is reached when the incremental cost of a further increase in safety precautions is greater than the incremental benefit (in increased safety) it will produce. Would contractualists calculate it differently from straight aggregationists?

Once again, the answer to that question is unclear. What it means to take 'reasonable precautions' is described in many different ways in the nonconsequentialist literature: You should behave in a fashion that is not negligent; that respects the legitimate interests of others to be free from harm; that treats others properly; that is not wrongful, unreasonable or unjustifiable; that gives

---

[63] That intuition, I think, underlies the view that acting in a fashion that is "sure" to result in harm is close to conclusive evidence that one acted without due care.

people what they are due.[64] But it is unclear whether the behavior demanded of us under any one of these formulations differs from the others, and whether any differs significantly from the optimal level of precaution dictated by aggregative techniques.

2.8. Limit the Ex Post POV to Serious Harms.

Another route that has been suggested to a middle way is to handle serious and nonserious harms differently. One approach, adopted by Scanlon, Keating and others, is basically a form of threshold deontology: When the potential harm passes some threshold of seriousness, those who (peeking ahead) learn they will be the big losers are given a veto or have the right to have their interests specially protected in some other fashion, but nonserious harms are handled by some form of aggregation.[65]

Like all disjunctive criteria, dividing the universe of potential harmful conduct in this fashion leaves contractualists with two unattractive alternatives. If a line is drawn and adhered to strictly, it produces results that are morally arbitrary. If one hour of pain is serious enough to be a serious harm, why not 59 minutes? On

---

[64] For a more detailed discussion of this literature, see Barbara H. Fried, "Is There a Coherent Alternative to Cost/Benefit Analysis?" (draft ms. 2010)

[65] Keating (2004). Keating proposes a hybrid system for deciding when to prohibit handling potentially harmful conduct, based on whether injuries are severe and irreparable. If they are, "it is essential that [we] induce appropriately great precaution….. Justice must be done at the time risk is imposed, by taking sufficiently stringent precautions." If they are not, the prohibition question should be resolved by straight CBA (permit if aggregate expected benefits exceed aggregate expected costs), relying on ex post compensation to take care of fairness concerns. Id. at 1876.

the other hand, if some flexibility is introduced to counter that arbitrariness, the distinction threatens to unravel because of transitivity issues.[66]

More significantly, the worst possible outcome of most risky activity is very bad—bad enough to be counted as serious by anyone's definition of serious. Thus, we are back at the impasse we face under the tiebreaker rule: If we measure the seriousness of potential harm by the worst possible outcome undiscounted by probabilities, as required by an ex post POV, then virtually all conduct threatens serious harms, and we are back to moral gridlock.

The second approach is some form of continuous function that weights expected harms more heavily, the more serious they are. While this approach is intended to produce different answers from conventional cost benefit analysis and other forms of unweighted aggregation, it is still a form of aggregation, in the sense that harms to me can be offset by benefits to you (albeit at a different ratio than under unweighted aggregation). It is also not clear what the justification is for giving more weight to serious harms than such harms would be given in unweighted aggregation. If the argument is that serious harms are, well, a lot more serious, that obvious truth should be reflected in the subjective disutilities that individuals assign to serious and trivial harms respectively. That is to say, a properly done cost/benefit calculus based on subjective preferences would normally assume steeply increasing marginal disutility as one moves up the scale of harms. If contractualists are evaluating expected outcomes based on subjective preferences as well, what is the argument for putting even more relative weight on serious

---

[66] For one version of the transitivity problem, see Parfit (2011, p. [ ]).

harms than the would-be victims themselves do?[67] If they are evaluating them based on something else, what is that something else? A paternalistic calculus of what we think people will prefer in the longrun? A perfectionist account of what they ought to prefer? A desire to protect society from ex post revolt against ex ante agreements that turn out very badly?[68] And how do we go about figuring out what weight serious harms 'deserve' under any of these criteria?

---

[67] It is not clear that proponents of this approach *are* weighting harms differently from welfarists. Consider in this regard the following hypothetical offered by Parfit. We have a fixed quantity of medicine. We could use all of it to increase Blue's lifespan from 30 to 70 years, or use it to increase some number of others' lifespan from 70 to 75 years. Parfit's intuition is that we should give it all to Blue, because Blue is worst off, as long as there are no more than (say) twenty of the others whose life we could increase by five years; but once the numbers of those others gets "very large"—a hundred or a million-- we should give it to them, because it would in Parfit's words be "too extreme" to ignore that aggregate benefit. (2011, pp. [ ]). Parfit describes this set of intuitions as a kind of weighted prioritarianism that "gives slightly more weight to the moral claims of people who are slightly worse off, and much more weight to the claims of people who are much worse of." (Id. at . But it is not clear that this is any different from the utilitarian solution, if we assume (plausibly) declining marginal utility of benefits in many realms. Suppose that the subjective value most people would place on increasing their lifespan from 30 to 70 is a lot more than eight times the subjective value they would place on increasing it from 70 to 75. (To put it another way, the marginal utility of one additional year of life decreases dramatically the longer one lives.) In that case, the straight utilitarian calculus, contra Parfit, is *not* indifferent to how the incremental years are distributed among different people, because the distribution affects their subjective value. A utilitarian calculus might plausibly conclude that we should give the forty years to Blue if the alternative is giving five years to twenty different people, but at a certain point, when the numbers of those other people got large enough, we should give the incremental years to those other people instead. In reaching that same conclusion, is Parfit just expressing that utilitarian calculus in different language, or is he doing something else?

[68] It is often unclear whether contractualists would have us value complaints about proposed principles based on individuals' presumed subjective preferences, some perfectionist notion of relative importance, or something else. I read the Scanlon of What We Owe to Each Other, in contrast to the Scanlon of "Preference and Urgency," to be working within subjective preferences for the most part. That is to say, in formulating our complaints about general principles, we are entitled to rank-order outcomes based on our own preferences, unless those preferences are so eccentric or offensive as to make it unreasonable to ask others to sacrifice for them. This last qualification explains why Scanlon can take it for granted that serious bodily harm and death, prima facie, trump other complaints: not because they generate objectively more urgent needs (although they do), but because we think almost everyone's subjective preferences would rank-order them ahead of any other "goods." But I'm not sure much turns on the answer to this for my purposes. Whatever the basis for rank-ordering outcomes, contractualists need to explain how they get to the conclusion that we should place relatively greater value on avoiding a serious harm than individuals themselves would.

2.9. Consider Every Case from Both Points of View

Concluding that a pure ex ante POV and a pure ex post POV are both dead-ends for contractualists, James Lenman suggests the solution is to meld the two perspectives, in a kind of reflective equilibrium. "If we think then of *ex ante* commitment and rejection as morally constrained and sensitive to the strains of commitment, and if we think of *ex post* agreement as reasonably sensitive to the earlier epistemic perspective in which principles have been applied in conditions of risk, it perhaps become easier to think of these approaches as complementary rather than necessarily competing."[69] Applying the method to Bernard Williams's famous hypothetical of the non-negligent lorry driver who, through no fault of his own, kills a young child who has darted into the street, Lenman argues we should accept *both* that the lorry driver's conduct must be justifiable to the parents from their ex post (heartsick) perspective and also that "[e]ven ex post there may be constraints on what [the parents] can reasonably reject." "If we take this point on board," Lenman concludes, "we can go some way along with Scanlon in privileging the perspective of the most burdened affected person ex post provided we also require that person to make due allowance for the epistemic perspective of the agent ex ante and it is here that probability information plausibly, properly, and indeed necessarily kicks in."[70]

---

[69] James Lenman (2008, p. 116). An ex post POV, Lenman argues, cannot "supply an action-guiding principle—that is, a principle that will tell the driver what to do. To the extent it has any influence on actions, it will "render contractualism excessively risk averse." On the other hand, an ex ante POV "threaten[s] to make contractualism in its upshot unattractively close to utilitarianism and ignores the reality that reasonable conduct with a bad outcome produces "strains of commitment." Id. at 115-117.

[70] Lenman (2008, pp. 115-116).

Is it possible to have it both ways? I don't think so. If by "due allowance," Lenman means that the parents can object ex post *only* if the driver's actions were unreasonable *and* that unreasonableness must be judged from an ex ante epistemic perspective, then the parents' ex post perspective does no work. We have a purely ex ante system in which the numbers are all that count. This seems to be Lenman's own conclusion in the lorry driver case: Considering the driver's conduct from the driver's own (ex ante) POV, Lenman asserts that the parents "should plausibly reasonably conclude the driver's choice not to be morally at fault."

If, on the other hand, by "due allowance" Lenman means that the parents' ex post objections will sometimes convert an action that was reasonable ex ante into one that in retrospect we judge to be unreasonable, then Lenman has recreated the paradox of moral luck that he rightly argues dooms a pure ex post POV as a theory of action. On what basis do we decide whether or not the parents' ex post objections are reasonable, if not on the basis of the ex ante reasonableness of the driver's conduct? If we decide the parents' objections are reasonable, do we therefore conclude, in retrospect, that the lorry driver should have acted differently? And what follows from that? That the parents may blame the lorry driver even though he acted just as we would have had him act? That they are entitled to compensation? In the end the only restriction Lenman himself appears to put on the ex ante POV concerns the rescission cases discussed above.

2.10. Adopt an Ex Post POV with a Strong Assumption of Altruism.

In his recent Tanner Lectures, Derek Parfit proposed to save Scanlonian contractualism from the more implausible implications of not letting the numbers count by accepting its commitment to count only 'individual reasons' grounded in rational motivations, but imputing to individuals a very substantial degree of altruism: "According to objective theories of the kind that I believe to be the truest and the best, we have strong reasons to care about our own well-being… But our own well-being is not, as Rational Egoists claim, the one supremely rational ultimate aim. We could rationally care as much about some other things, such as the well-being of others."[71] By assuming we do in fact care as much about others as about ourselves, Parfit gets the numbers to count—not, as in aggregative reasoning, because an impartial conception of justice requires it, but because everyone will individually choose to have them count, even when doing so cuts against their rational self-interest. Thus, taking up Taurek's famous dilemma about whether to save one person on one rock or five people on another, Parfit concludes we can get to the right result (save the five) without departing from Scanlon's "individual reasons" restriction, because the one, being altruistic, could rationally choose that the five people be saved in place of himself.[72]

As critics have noted, Parfit's implausibly strong assumption of altruism seems to smuggle in his own normative preference for the aggregative

---

[71] Parfit (2011, p. [ ]).
[72] Parfit (2011, p. [ ]).

('optimific') solution through the back door. But for present purposes, I am less interested in whether Parfit succeeds in giving a nonconsequentialist justification for his conclusion than in what that conclusion is—in particular, whether it offers a genuine alternative to aggregation.

Parfit suggests it does, in part because he does presume a limit to our altruism,[73] in part because his version of consequentialism is more ecumenical than welfarism, taking account of preferences about the distribution of burdens and benefits and preferences for autonomy over well-being, at least in certain contexts.[74] But it remains to be seen whether any of these differences will, in Parfit's hands, produce outcomes that deviate significantly from conventional aggregation. There is reason to be skeptical here, I think. Consider in this regard Parfit's handling of what he calls the Harmful Means Principle. On its face, the principle appears to endorse the nonconsequentialist position that how consequences are brought about may matter, independent of the consequences themselves, at least in some cases (e.g., whether it is permissible to throw one person onto the trolley tracks to save five). But here is how Parfit elaborates what the Harmful Means Principle requires of us: It is wrong to harm someone *as a means of* achieving some aim unless "(i) there is no better way to achieve this aim, and (2) given the goodness of his aim, the harm we impose is not

---

[73] Parfit (2011, p. [ ])(rejoinder to Susan Wolf).

[74] Welfarists would also give such preferences independent weight in measuring how well things go (the consequences), counting our "taste" for autonomy just as it counts our "taste" for material consumption or health.) Whether Parfit would weight these preferences differently from welfarists is unclear.

disproportionate, or too great."[75] Once again, we are left with the question how, if at all, this differs from how we would assess options in a conventional optimific calculus?

3. Facing the Inevitability of Tradeoffs

In my view, the only clearly viable and meaningful alternatives to conventional aggregation that contractualists have proposed to date are the distributionally sensitive versions of an ex ante POV discussed in 3.2 and 3.3 above. It is possible that some of the other approaches, once fleshed out, will offer a clear alternative to aggregation. I hope this article, if it accomplishes nothing else, will encourage contractualists to turn their attention to that task.

I am doubtful about the ultimate success of those efforts, however, for two reasons. First, even in their very sketchy form, the hybrid approaches offered to date suggest exactly the sorts of interpersonal tradeoffs that are at the heart of standard aggregation. Second, and more importantly, the problem on the table—defining the permissible limits of socially useful but risky conduct—is one that, in my view, cannot be solved without such tradeoffs. The reason for this is straightforward. Scanlonian contractualism, like most Kantian-inflected versions of nonconsequentialism, is committed to two foundational principles: that each of us has a legitimate interest in pursuing our own projects in life (freedom of conduct), and in doing so, each of us has a right not to be interfered with by others, most especially by serious bodily harm (freedom from harm inflicted by others).[76] But

---

[75] Parfit (2011, pp. [ ]).

[76] For one representative statement, see Scanlon (2000, p. 313): "[T]he idea that each person has a

most conduct in pursuit of one's own projects carries with it some risk (however remote) of harming others. As a result, we cannot protect one of those two interests without compromising the other.

Where a given act is very, very unlikely to harm others (changing channels on the TV, making the bed, answering a letter), we can effectively ignore the risk associated with it. But such conduct, precisely because it poses no meaningful threat to others, requires no restriction under any normative criteria. Harm that results from criminal conduct is an easy case in the opposite direction: Under any normative criteria, we place no value on leaving the criminal free to pursue his own (criminal) projects, and hence we need not compromise the would-be victims' right to be free from harm in order to protect it.

In between those two extremes lies the vast universe of conduct that is socially productive and hence prima facie permissible, but that poses some meaningful risk of harm to others. The inherent conflict between freedom of conduct and freedom from harm is posed in stark and unavoidable terms in such garden-variety cases of risk. The implicit compromise struck is expressed in lots of different ways (you can't harm someone else "negligently" or "unreasonably," the right to be free from harm must give way when the social costs of adhering to it are just too great). These different formulations may or may not strike that compromise differently from each other or from unweighted aggregation. But in the end they all (necessarily) reduce to a tradeoff between these two interests.

---

special claim to and authority over his or her own life and body" is "one that no one could reasonably reject."

However unattractive one might find aggregation in principle and however unadministrable it might be in practice, it at least faces that necessity squarely. In my view, nonaggregationists have yet to do so, or explain how it can be avoided.

References

Ashford, Elizabeth. 2003. The Demandingness of Scanlon's Contractualism. *Ethics* 113: 273-302.

Brand-Ballard, Jeffrey. 2004. Contractualism and Deontic Restriction. *Ethics* 114:269-300.

Brock, Dan. 1998. Aggregating Costs and Benefits. *Philosophy and Phenomenological Research*. 51: 963-67.

Broome, John. 1978. Trying to value a life. *Journal of Public Economics* 9:91–100.

Fletcher, George P. 1972. Fairness and Utility in Tort Theory. *Harvard Law Review* 85:537-573.

Fletcher, George P. 1985. The Search for Synthesis in Tort Theory. *Law and Philosophy* 2:63-88.

Hansson, Sven Ove. 2007. Risk and Ethics. In *Risk: Philosophical Perspectives*, ed. Tim Lewens, 21-35. London: Routledge.

Heinzerling, Lisa. 2006. Knowing Killing and Environmental Law. *N.Y.U. Environmental Law Journal* 14:521-534.

Herman, Barbara. 1996. *The Practice of Moral Judgment.* Cambridge, MA: Harvard University Press.

Keating, Gregory. 2004. Rawlsian Fairness and Regime Choice in the Law of Accidents. *Fordham Law Review* 72:1857-1922.

Kamm, F.M. 1993. *Morality, Mortality, Volume I: Death and Whom to Save From It.* Oxford University Press.

Kamm, F.M. 2007. *Intricate Ethics.* Cambridge: Harvard University Press.

Kavka, Gregory. 1979. The Numbers Should Count. *Philosophical Studies* 36:285-294.

Kumar, Rahul. 2003. Who Can Be Wronged? *Philosophy & Public Affairs* 31:99-118.

Kumar, Rahul. 1999. Defending the Moral Moderate: Contractualism and Common Sense. *Philosophy and Public Affairs* 28:275–309.

Lenman, James. 2008. Contractualism and Risk Imposition. *Politics, Philosophy and Economics* 7:99-122.

Miller, Sarah. 2006. Need, Care and Obligation. *Royal Institute of Philosophy Supplement* 80:137-160.

Munoz-Dardé, Véronique. 2009. Conversations with an Amish Farmer: Risk and Reasonable Rejection. URL:

Otsuka, Michael. 2000. Scanlon and the Claims of the Many Versus the One. *Analysis* 60: 288-93.

Parfit, Derek. 1978. Innumerate Ethics. *Philosophy and Public Affairs* 7:285-301.

Parfit, Derek. 2011. *On What Matters*. Forthcoming, Oxford University Press.

Parfit, Derek. 2003. Justifiability to Each Person. *Ratio* XVI:368-90.

Perry, Stephen. 2001. Responsibility for Outcomes, Risks, and the Law of Torts. In Gerald Postema, ed., *Philosophy and the Law of Torts*, 72-130.

Raz, Joseph. 2003. Numbers, With and Without Contractualism. *Ratio* XVI:346-67.

Rawls, John. 1971. *A Theory of Justice.* Cambridge, MA: Harvard University Press.

Rawls, John. 1999. *A Theory of Justice (revised ed.).* Cambridge, MA: Harvard University Press.

Reibetanz, Sophia. 1998. Contractualism and Aggregation. *Ethics* 108:296-311.

Ripstein, Arthur. 2006. Beyond the Harm Principle. *Philosophy and Public Affairs* 34:216-246.

Ripstein, Arthur. 2009. *Force and Freedom*. Cambridge, MA: Harvard University Press.

Sanders, John T. 1988. Why the Numbers Should Sometimes Count. *Philosophy and Public Affairs* 17:3-14.

Scanlon, Thomas. 1975. Preference and Urgency. *Journal of Philosophy* 72:655-669.

Scanlon, Thomas. 1982. Contractualism and Utilitarianism. In Amartya Sen and Bernard Williams, eds. *Utilitarianism and Beyond*, 103-28.

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Scanlon, Thomas. 2000. Intention and Permissibility. *Aristotelian Society* 74:301-317.

Scheffler, Samuel. 1994. *The Rejection of Consequentialism*. Oxford University Press.

Taurek, John. 1977. Should the Numbers Count? *Philosophy and Public Affairs* 6:293-316.

Williams, Bernard. 1973. *Utilitarianism: For and Against.* With J.J.C. Smart. Cambridge: Cambridge University Press.