Fixing Aggregation

Abstract: How should we distribute harm when harm is unavoidable? Many theorists reject the idea that we should simply minimize the aggregate "weight" of harm. But they also reject the idea that we should simply avert the weightiest individual harm(s). They instead endorse *Moderate Aggregation* — the view that some, but not all, harms of inequivalent weight trade-off against one another. On this view, there may be some number of people threatened with a broken finger that should be rescued over one person threatened with a broken arm; and yet there may be no number of people threatened with a mild papercut that should be rescued over one person facing the loss of their legs. I argue that even the most sophisticated versions of this theory currently on offer suffer a common, fatal flaw. This flaw can be traced to the fact that these theories attempt to explain the moderate nature of harm aggregation in terms of differences in harm *degree*. I argue that defenders of moderate aggregation would do much better to appeal to differences in harm *kind*.

1 Introduction

Sometimes harm is unavoidable. Sometimes — as frequently happens in war or medicine or disaster relief — we can only choose how harms are distributed. We can only save some.

In such cases, how should we choose between distributions of harm that vary only with respect to the *quantity* of persons who suffer harm and the *severity* of those harms?[1]

One possible answer is

[1]There are of course other factors that make a difference as to who we should save. For example: our relationship to the victims, our responsibility for the harms in question, the responsibility of others for the harms in question, whether the victims want to be rescued, whether any of the victims are liable to be harmed, whether we would be "doing" or "allowing" harm by failing to rescue. For the purposes of this paper we will hold these other factors fixed in each case, so as to focus in on the effects of the quantity and quality of harms.

> **No Aggregation**. All else equal, a distribution D_1 is more choiceworthy than a distribution D_2 just in case D2 contains a *more serious harm* than D1.[2]

This answer has very counterintuitive implications. Consider:

> *Legs.* You must choose between saving (i) one person from the loss of both her legs, or (ii) saving each of one billion people from the loss of one of their legs.

No Aggregation tells us that you should choose option (i), since the loss of two legs is a greater harm than the loss of a single leg. But this seems false. Intuitively, you should choose option (ii), saving the many over the one.

To get this result our theory must allow for weaker harms to (at least sometimes) collectively win out over stronger harms. Our theory must allow the choiceworthiness of a harm distribution to be a function of the *aggregate* weight of certain harms in the distribution. But the aggregate weight of *which* harms? The most straightforward answer is: All of them. Call this approach

> **Pure Aggregation**. All else equal, a distribution D_1 is more choiceworthy than a distribution D_2 just in case D_1 contains less aggregate harm than D_2.[3]

But such unrestricted aggregation also has counterintuitive implications. Consider:

> *Papercuts and Amputation.* You must choose between saving (i) each of one billion people from a minor papercut, or (ii) saving one person from the loss of both her legs.

Intuitively, you should save the one person from the loss of her legs, no matter how many papercuts you could prevent instead. But Pure says otherwise. On the plausible assumption that the weights of both sorts of harm are non-zero

[2] See Anscombe (1967) and Tuarek (1977) for a defense of No Aggregation (or something very close to it). See also: Munoz-Dardé (2005), Thomas (2012), and Doggett (2013).

[3] See Halstead (2016), Horton (2017, 2018, 2020), and Norcross (1997, 2002) for a defense of Pure Aggregation (or something very close to it). It should be noted that some theorists formulate their theories in terms of a different "currency" of aggregation. Some prefer, for example, to aggregate *reasons* or *claims* or *complaints*, rather than *harms*. The important lessons of this paper will apply equally well to theories framed in these other currencies. I focus on the aggregation of *harm* because it is simpler to do so. Once we focus on reasons more generally, for example, we need to take account of both costs and benefits, and this raises complications about the "asymmetry" of costs and benefits (on this topic, see Shiffrin (2012)).

and finite, there is *some* number of papercut harms such that their aggregate weight is greater than the weight of the harm of losing on's legs.[4]

The counterintuitive implications of Non-Aggregation and Pure have led many theorists to seek out a middle way between these two extremes. The goal has been to locate a theory that allows *some, but not all*, inequivalent harms to trade-off against one another. Let's call any theory with this feature a *moderately aggregative* theory.

Moderately aggregative theories must draw a line somewhere. Structurally, moderate theories are committed to

> *Lexical Tiers.* In any given choice situation, harms can be sorted into at least two "tiers" such that the agent should minimize any aggregate weight of harm from any higher tier over any aggregate weight of harm from any lower tier.

The most straightforward way to think of such tiers are as "fixed" or "invariant" categories such that harms of a given severity belong to the same tier across all possible choice situations. On this approach, we should save one person's legs over saving any number of people from papercuts because the harm of losing one's legs is a *type* of harm that belongs to a "higher" class than the harm of suffering a papercut.

This *Fixed* view, however, has not been taken seriously by moral philosophers. It's not hard to see why. Moral philosophers often make the simplifying assumption that harms are measurable on a *single* ratio scale — a scale that allows us to talk meaningfully about ratios of differences in harm (e.g., "The difference in severity between harm x and harm y is twice as great as the difference in severity between harm y and harm z.") Given this assumption, it seems we should think of fixed tiers as *non-overlapping, jointly exhaustive subregions* of a single harm-severity scale. And this seems implausible. Why would the harm-severity scale partition into subregions such that higher subregions *lexically dominate* lower subregions?

[4] Defenders of Pure should find both of these assumptions extremely plausible. Here's a reason the defender of Pure should accept that the weight of the harm of losing two legs is finite: there are surely *worse* harms out there. Here's another reason: if such a harm is infinitely bad, then Pure would implausibly predict that rescuing one person from such a harm is just as choiceworthy as rescuing one-billion people from such a harm. And here's a reason the defender of Pure should accept that the weight of a papercut harm is non-zero: if it weren't, then Pure would implausibly predict that rescuing one person from a papercut is just as choiceworthy as rescuing one-billion people from a papercut.

Why would merely *quantitative* differences make such a difference? What would be so special about the degree of severity at which one tier borders another such that harms only *slightly more* severe take lexical priority over harms only *slightly less* severe?

Worries like these have led Fixed to be virtually ignored in the "harm aggregation" literature. Proponents of moderate aggregation have opted instead for "variable" tiers. On this view, harms of severity s can fall into *different* tiers in different choice situations. What determines the tier to which a given harm belongs isn't the absolute severity of that harm, but rather how great the *difference* in severity is between that harm and certain other harms at play in the choice situation.[5]

By way of illustration, consider one of the simpler versions of the view — a version suggested by Kamm (1998, 2005, 2007). On Kamm's view, all of the harms a person can avert in a given choice situation can be sorted into one of two tiers — the "relevant" and the "irrelevant" (2007: 297-298, 484-486).[6] The most weighty harm that can be averted *in that situation* falls into the relevant tier. So too do any harms that are "close" in weight to that most severe harm. The harms that are not close in weight to the most severe harm fall into the irrelevant tier.[7] On this theory, possible papercuts are irrelevant in the presence of possible leg loss because a papercut harm is not "close enough" in weight to the harm of losing one's legs. But the possible loss of one leg is relevant in the presence of the possible loss of two legs because the harm of losing one leg *is* close in weight to the harm of losing two legs.

As many critics have noted, however, simpler versions of Variable have absurd implications of their own. They violate a number of very plausible "consistency" principles.[8] In response to these challenges, defenders of Variable have, in recent years, developed increasingly complex versions of the view and argued that these more complex versions avoid the most damning objections.

In the first half of this paper I argue that this isn't so. The most sophisticated

[5] This approach can be traced back to the late 1990s, when it was proposed by both Kamm (1998: 297-298, 484-486) and Scanlon (1998:238-241).

[6] Kamm is one of the aforementioned theorists who prefers to frame her theory in terms of the aggregation of *reasons*, rather than *harms*. See footnote 3.

[7] The *closeness-in-weight* relation employed by defenders of Variable is non-transitive.

[8] The most damning of these objections is developed by Tomlin (2016). See also Halstead (2016). There are other objections that have been leveled against Variable as well, most notably the objection that it cannot handle cases involving "uncertainty". I'll discuss this objection in §5.3.

versions of Variable are just as damned as the least. In §2 and §3 I make the case that *every* extant version of Variable violates an extremely compelling moral principle that I call *More Ain't Less*.

This failure, as we'll see, is unsurprising. On the Variable view, whether harms of a given severity are "relevant" depends on what other harms are in play; tier membership is situation-relative. In some choice situations papercut harms are irrelevant (e.g., in the presence of possible leg loss); in other situations papercut harms are relevant (e.g., when there are no other harms that can be averted). The variability of the Variable view guarantees that the choiceworthiness ordering over harm distributions is non-transitive, and it is to be expected that non-transitive orderings will violate all sorts of "consistency" principles.[9]

One option for proponents of moderate aggregation is to bite this very bitter bullet. In §4 and §5 I argue that a much better option is to reconsider the Fixed view — a view that delivers the consistencies that Variable cannot. Fixed deserves attention, not just because it hasn't received any, but because it is a good deal *more* plausible than the Variable view. Or so I argue.

2 Variable & More Ain't Less

Every version of Variable on the market agrees on the following three claims:

> *Two Tiers.* In any given choice situation harms can be sorted into two tiers — the "relevant" and the "irrelevant".
>
> *Relevance by Closeness.* Whether a harm is relevant in a given choice situation depends on the *closeness in weight* between that harm and certain other harms in that choice situation.
>
> *Minimize Relevant Harm.* All else equal, a distribution D_1 is more choiceworthy than a distribution D_2 just in case D1 contains less aggregate *relevant* harm than D2.[10]

[9] The choiceworthiness ordering is non-transitive on the Variable view in that the view is inconsistent with TRANSITIVITY: *If you should choose harm distribution A over B, and you should choose distribution B over C, then you should choose A over C.*

[10] Things are actually a bit more complicated than this, in two respects. First, some defenders of Variable allow irrelevant harms to "break ties": they claim that if two distributions do equally well at minimizing relevant harm, then the distribution that contains less irrelevant harm is more choiceworthy. Second, some defenders of Variable would accept Minimize Relevant Harm in cases where a choice situation consists of only two harm distribution options, but not in situations with three or more options. I ignore both of these complications. My objections to

A "closeness" measure plays a central role in this theory. The closeness measure is what sorts harms into the relevant and irrelevant tiers. If a harm is "close" in weight to its "anchoring" harm, it is relevant; otherwise it is irrelevant. This much all proponents of Variable agree on. Where different versions of Variable come apart is with respect to the details of the anchoring relation. Here, for example, are two of the more digestible anchoring rules that have been proposed in the literature:

> *The Simple Rule.*[11] Every harm is anchored to the *weightiest harm in the choice situation* (i.e., the weightiest harm the agent can avert). Harms that are close in weight to that harm are relevant. Harms that are not close in weight to that harm are irrelevant.

> *The Competition Rule.*[12] There is not one harm to which all harms in a choice situation are anchored. A harm is anchored to the weightiest harm with which it "competes" in that choice situation (i.e., the weightiest harm with which it cannot be jointly averted). Harms that are close in weight to their anchoring harm are relevant. Harms that are not close in weight to their anchoring harm are irrelevant.[13]

Many proponents of Variable have rejected these two rules on the grounds that they have absurd implications when we consider *series* of cases. The most devestating criticism of this sort comes from Tomlin (2018). In response, defenders of Variable — Tomlin included — have proposed more sophisticated variants of the view. In what follows I argue that absurd results in series of cases

---

Variable do not turn on whether irrelevant harms break ties or not, and my objections apply even in simple two-option cases.

[11] See Kamm (2007: 297-298, 484-486). This is perhaps also Scanlon's (1998: 239-240) view.

[12] See Voorhoeve (2014: 66–67).

[13] Here is an unnoticed problem for anchoring rules that appeal to this notion of *competition*. Suppose there is one person, X, in danger of permanent paralysis and one billion Ys each in danger of suffering a papercut. You can (i) rescue one half of the papercut victims, (ii) rescue the other half of the papercut victims, or (iii) rescue any one pairing of one papercut victim with the paralysis victim (i.e., you can rescue X and Y_1, or X and Y_2, or ...). In this scenario, every papercut harm *can* be jointly averted with the paralysis harm. The weightiest harms with which each of the papercut harms *cannot* be jointly averted, then, are other papercut harms. And since papercut harms are close in weight to one another (being identical in weight), rules like Competition are forced to say that all of the papercut harms in this situation are *relevant*. But this means that, given how many papercut harms are at stake, you should rescue half of the Ys from papercuts over rescuing one person from paralysis and one person from a papercut. This flies in the face of the moderate project. It was implications like this that motivated the rejection of Pure Aggregation and the search for a theory of moderate aggregation in the first place! Competition-based anchoring rules need a different notion of competition if they are to deliver the results moderate theories are intended to deliver.

are not limited to simple versions of Variable. *Every* version of Variable has this result. Specifically, I argue that every version of Variable violates the following (extremely plausible) principle in rather shocking ways:

> MORE AIN'T LESS. If you should choose harm distribution A over B, then you should choose A^+ over B^+, where A^+ is the result of adding n-many harms of weight w to A, B^+ is the result of adding m-many harms of weight w to B, and $n > m$.[14]

Before we get to the more sophisticated versions of Variable, it will help to start by considering how simpler versions (e.g., the Simple and Competition rules) violate this principle. Let's suppose we have three types of harm — severe, moderate, and mild — that satisfy the following conditions:

- Moderate harms are close in weight to severe harms.
- Mild harms are close in weight to moderate harms.
- Mild harms are not close in weight to severe harms.
- The weight of 1 severe harm = the aggregate weight of 10 moderate harms = the aggregate weight of 100 mild harms.

Now imagine you face the following choice:

> *Fire 1.* Two buildings are ablaze. You can save the persons trapped in building A or the persons trapped in building B, but you cannot get to both buildings. Building A contains 20 people in danger of suffering moderate harm. Building B contains 1,000,000 people in danger of suffering mild harm.

By stipulation, mild harms are close in weight to moderate harms. In this situation, then, the Simple Rule identifies all of the harms as relevant. Since Variable would have you minimize aggregate relevant harm, and since 1,000,000 mild harms outweighs 20 moderate harms, you should should avert the mild harms — i.e., rescue the persons trapped in building B.

But now suppose that:

> *Fire 2.* After contemplating your options in Fire 1 you decide to

[14] Tomlin argues that the Simple version violates what he calls the *Principle of Addition*, and that the Competition version violates what he calls *Equal Consideration for Equal Claims*. My principle is somewhat different, but the cases I discuss in this section have the same structure as Tomlin's.

> rescue the persons trapped in building B. On your way to the building, the situation changes. You come to learn that another person has fallen into danger. Building B now contains 1,000,000 people in danger of suffering mild harm *and* 1 person in danger of suffering severe harm. The status of those in building A remains as before.

	Building A		Building B	
Fire 1	moderate	x20	**mild**	**x1,000,000**
Fire 2	**moderate**	**x20**	mild	x1,000,000
			+ severe	x1

(The bold font indicates the group that the Simple Rule directs you to save.)

Surely the fact that another potential victim has been added to building B gives you *all the more reason* to rescue those trapped in that building. But the Simple Rule absurdly implies otherwise. It implies that the addition of the new victim demands that you change course and rescue those trapped in building A. It has this result because the presence of a severe harm renders all the mild harms *irrelevant*, and the aggregate weight of 20 moderate harms is greater than the weight of 1 severe harm. Thus we get a violation of *More Ain't Less*.

The Competition Rule delivers a consistent verdict in Fire 1 and Fire 2, since it counts the mild harms as relevant in both cases. But the Competition rule violates *More Ain't Less* in other cases. Consider:

> *Fire 3.* Two buildings are ablaze. You can save the persons trapped in building A or the persons trapped in building B, but you cannot get to both buildings. Building A contains 1 people in danger of suffering severe harm. Building B contains 20 people in danger of suffering moderate harm.

The Competition Rule tells you to rescue the persons trapped in building B. According to this rule, we don't compare each harm with the strongest overall harm, but with the strongest harm *with which it competes*. In this case, the severe harm in A must be compared with the moderate harms in B, and vice versa. Since these harms are close in weight to one another, all of the harms in play are relevant. Since Variable would have you minimize aggregate relevant harm, and since 20 moderate harms outweigh 1 severe harm, you should rescue the persons trapped in building B.

But now consider:

> *Fire 4.* After contemplating your options in Fire 3 you decide to rescue the persons trapped in building B. On your way to the building, the situation changes. You come to learn that a great many more people have fallen into danger. Building A now contains the 1 person facing severe harm and 1,000 other people facing mild harm. Building B now contains the 20 people facing moderate harm and 1,000,000 other people facing mild harm.

	Building A		Building B	
Fire 3	severe	x1	**moderate**	**x20**
Fire 4	**severe**	**x1**	moderate	x20
	+ mild	**x1,000**	+ mild	x1,000,000

(The bold font indicates the group the Competition Rule directs you to save.)

People in danger of suffering mild harm have been added to both buildings. But a great many more of them have been added to building B. Surely this gives you *all the more reason* to rescue those trapped in building B. But the Competition Rule absurdly implies otherwise. It implies that the addition of the new victims demands that you change course and rescue those trapped in building A — in violation of *More Ain't Less*. This is because the Competition Rule anchors harms to their *competition*. This forces us to treat the two sets of mild harms differently. The rule must count the 1,000 mild harms that are added to building A as relevant, but the 1,000,000 mild harms that are added to building B as irrelevant. Since the aggregate weight of 1,000 mild harms + 1 severe harm is greater than the aggregate weight of 20 moderate harms, the Competition Rule would have you rescue the victims in building A. Thus we get a violation of *More Ain't Less*.

3 The Balancing Approach

The problem we've thus far encountered is this. If Variable is to deliver a consistent verdict in both pairs of Fire cases, it can't let the presence of a severe harm *entirely* neutralize the deontic impact of mild harms in cases like Fire 2 and Fire 4. At the same time, of course, the theory can't allow mild harms to

"trade off" with severe harms, or else the theory would fail to be a moderate theory at all. What the proponent of Variable needs, it seems, is a way for mild harms to make *some* kind of a difference when they compete with severe harms, without letting mild harms trade off against severe harms.

A recent innovation proposes such a way. On what I'll call the Balancing approach, the relevant/irrelevant sorting occurs only on those harms that "survive" a *sequential cancellation procedure*.[15] The details of these views are (quite!) complicated, but the core idea behind the use of a cancellation procedure is that we should *treat like harms alike*: if a set of harms in one group is of the same aggregate weight as a set of (individually close) harms in a competing group, then the two sets of harm are balanced, and they cancel each other out; we can effectively ignore them. So before we do any relevant/irrelevant sorting, we "cross off" those sets of close and competing harms that are equal in weight to one another. Severe harms in one group can balance with (an equal number of) severe harms in the other group; severe harms can balance with (a larger number of) moderate harms; moderate harms can balance with (an equal number of) moderate harms; moderate harms can balance with (a larger number of) mild harms; and mild harms can balance with (an equal number of) mild harms.

Here's an illustration of the idea. Let's continue to suppose that

- Moderate harms are close in weight to severe harms.
- Mild harms are close in weight to moderate harms.
- Mild harms are not close in weight to severe harms.
- The weight of 1 severe harm = the aggregate weight of 10 moderate harms = the aggregate weight of 100 mild harms.

Now consider a choice between rescuing group A or group B. A consists of 1 person facing severe harm. B consists of 10 people facing moderate harm and 100 people facing mild harm.

Group A		Group B	
severe	x1	moderate	x10
		+ mild	x100

[15] See Steuwer (2021), Tadros (2019), and van Gils & Tomlin (2020) for different versions of the Balancing Approach.

One sort of cancellation procedure is a "top-down" one. A "top-down" cancellation procedure would have us start at the top, with the most weighty harm in the situation. In this case, the most weighty harm is the severe harm in group A. We look to see if this harm balances with any close harms in group B. The severe harm in A cannot be balanced with the mild harms in B, since only harms that are close in weight can be balanced against one another. But the severe harm in A is perfectly balanced by the 10 moderate harms in B. So those harms cancel each other out. We cross them off, leaving us with only the 100 mild harms in group B.

Group A		Group B	
~~severe~~	~~x1~~	~~moderate~~	~~x10~~
		+ mild	x100

Since the remaining harms cannot be balanced with anything, the cancellation procedure comes to an end. Only now do we sort harms as relevant or irrelevant. Since the sorting ignores cancelled harms, the mild harms count as relevant in this situation. So you should rescue group B.

The addition of a cancellation procedure looks to give us what we were looking for: a way for mild harms to make a moral difference when they compete with severe harms, without letting mild harms trade off against severe harms. This feature of Balancing enables to theory to deliver a consistent result in both Fire 1/Fire 2 and Fire 3/Fire 4 (for the sake of space, I won't take the reader through the details of how). That's a very good result, and it has led some proponents of Variable to conclude that Balancing obeys the most plausible consistency principles that are violated by non-Balancing versions of Variable.

It doesn't. Balancing, like other versions of Variable, violates *More Ain't Less*. Consider:

> *Fire 5.* Building A contains 1 person in danger of suffering severe harm. Building B contains 1,000,000 people in danger of suffering mild harm.

Balancing, like all versions of Variable, tells you to rescue the person in building A. Mild harms are *not* close in weight to severe harms, and thus mild harms cannot be balanced with severe harms. So there is no cancelling to be done in

this case. We jump straight to the relevant/irrelevant sorting. Since the severe harm is the only harm that is relevant (on any anchoring rule), that's the harm you must avert.

Now consider:

> *Fire 6.* After contemplating your options in Fire 5 you decide to rescue the persons trapped in building A. On your way to the building, the situation changes. You come to learn that additional people have fallen into danger. 100 people facing moderate harm have been added to building A (alongside the 1 person facing severe harm); 20 people facing moderate harm have been added to building B (alongside the 1,000,000 people facing mild harm).

The move from Fire 5 to Fire 6 is analogous to the move from Fire 3 to Fire 4. We add harms of a single type to both groups, except that we added many *more* of these harms to the group that was previously the group you should rescue. Surely this gives you *all the more reason* to rescue that group. But Balancing absurdly implies otherwise. In Fire 5 Balancing tells you to rescue those in building A. In Fire 6, by contrast, Balancing tells you to rescue those in building B. It yields this verdict regardless of the order in which we balance harms against each other.

	Building A		Building B	
Fire 5	**severe**	**x1**	mild	x1,000,000
Fire 6	severe	x1	**mild**	**x1,000,000**
	+ moderate	x100	**+ moderate**	**x20**

If we use a "top-down" cancellation procedure, we start with the weightiest harm in the choice situation. In this case, the weightiest harm is the 1 severe harm in A. We look to see if this severe harm balances with any harms in B. It does. It balances with (and only with) 10 moderate harms in B. So these harms cancel each other out, leaving:

	Building A		Building B	
Fire 6	~~severe~~	~~x1~~	~~moderate~~	~~x10~~
	+ moderate	x100	+ moderate	x10

	Building A		Building B	
			+ mild	x1,000,000

The top-down procedure now has us look to the weightiest remaining harms — the moderate harms. The 10 moderate harms in B balance with 10 moderate harm in A, cancelling each other out:

	Building A		Building B	
Fire 6	~~severe~~	~~x1~~	~~moderate~~	~~x10~~
	~~+ moderate~~	~~x10~~	~~+ moderate~~	~~x10~~
	+ moderate	x 90	+ mild	x1,000,000

Next, the 90 moderate harms remaing in A balance with 900 mild harms in B, cancelling each other out. This leaves us with only a set of mild harms in B:

	Building A		Building B	
Fire 6	~~severe~~	~~x1~~	~~moderate~~	~~x10~~
	~~+ moderate~~	~~x10~~	~~+ moderate~~	~~x10~~
	~~+ moderate~~	~~x 90~~	~~+ mild~~	~~x900~~
			+ **mild**	**x999,100**

Given a top-down cancellation procedure, Balancing tells us that we should rescue group B. And thus we arrive at another very counterintuitive violation of *More Ain't Less*.

There are other possible cancellation procedures. And although our choice of cancellation procedure make a difference in some cases, it makes no difference here. Balancing's inability to obey *More Ain't Less* does not turn on the cancellation procedure we use. We arrive at the same problematic conclusion if we instead use, for example, a "bottom-up" procedure — a procedure that has us start with the weakest harms in play, and cancel "upwards" from there.

In Fire 5, recall, our choice of cancellation procedure makes no difference, since there are no competing harms that are relevant to one another. So if our choice

of cancellation procedure makes a difference, it must make a difference with respect to Fire 6. But consider how a bottom-up procedure would apply to this case. In Fire 6, the weakest harms in play are the mild harms in building B. So we start there. These harms can only balance with the 100 moderate harms in A. So 1,000 mild harms in B cancel out with the 100 moderate harms in A:

	Building A		Building B	
Fire 6	~~moderate~~	~~x100~~	~~mild~~	~~x1,000~~
	severe	x1	moderate	x20
			+ mild	x999,000

The remaining mild harms in B are not close in weight to any of their competitors, so they can't be matched with anything else. If we stop the procedure there, the result is that we should rescue group B (since 20 moderate harms outweigh 1 severe harm). If, on the other hand, we continue the procedure by looking to the *next* weakest harms, then we look to the 20 moderate harms in B. 10 of these cancel out with the 1 severe harm in A. This leaves behind only harms in B:

	Building A		Building B	
Fire 6	~~moderate~~	~~x100~~	~~mild~~	~~x1,000~~
	~~severe~~	~~x1~~	~~moderate~~	~~x10~~
			moderate	**x10**
			+ mild	**x999,000**

Once again we arrive at the problematic conclusion that we should rescue group B.[16]

Upshot: The only way to secure the verdict that we should rescue group A in *both* Fire 5 and Fire 6 is to prevent the 1,000,000 mild harms in building B from having *any* "cancelling" force. But permitting mild harms to have cancelling force when they compete with severe harms was the very feature that enabled Balancing to deliver consistent results in Fire 1/Fire 2 and Fire 3/Fire 4. Giving up this feature would lead us right back to those counterexamples. Defenders

[16] Notice that although the bottom-up procedure ultimately yields the same verdict as the top-down procedure, the two procedures do not leave behind exactly the same harms in building B.

of Variable are, therefore, trapped in a dilemma. If a theory allows mild harms to have cancelling force when they compete with severe harms, then the theory delivers inconsistent results in Fire 5/Fire 6. If the theory does not allow mild harms to have cancelling force when they compete with severe harms, then the theory delivers inconsistent results in Fire 1/Fire 2 or Fire 3/Fire 4. Either way, a violation of *More Ain't Less*. The Balancing approach, despite all its complexity and ingenuity, does not help Variable to obey this moral principle.

4 The Fixed View & More Ain't Less

I have thus far argued that every extant version of Variable violates *More Ain't Less*. This shouldn't be too surprising. Bizarre logical features are to be expected from a theory that delivers a *non-transitive* choiceworthiness ordering over harm distributions.

And of course the reason Variable delivers a non-transitive choiceworthiness ordering is right there in the name: it is the use of *variable* tiers. The fact that the "tier membership" of harms of a certain severity can change from choice situation to choice situation is what produces a non-transitive ordering, which is what in turn produces the violation of *More Ain't Less*.

The obvious way to avoid these results is to go in for *fixed* tiers. There are two ways to flesh out this idea. The first we might call the *monoscalar* interpretation. In §1 we noted that moral philosophers often make the simplifying assumption that the severity of a harm is best represented as a real number on a *single* ratio scale. Given this assumption, it makes sense to think of tiers as *subregions* of this single harm-severity scale. Geometrically, the idea is this: we represent the weight of a harm as a point on a number line (the harm-severity scale); we represent various tiers as non-overlapping, jointly exhaustive subregions of that line; and we say that a harm "belongs" to a tier just in case it is located within that subregion of the scale.

But we needn't treat a simplifying assumption as a data point. Perhaps harms are not measurable along only one dimension. Perhaps there are many dimensions to a person's well-being, and a person's well-being can be measured along each of these dimensions. If so, then the measure of a harm may be best represented, not as a single number, but as an ordered n-tuple $\langle n, \ldots, m \rangle$, where n represents the measure of the harm along the most important dimension, and m represents the

measure of the harm along the least important dimension. On this interpretation — call it the *multiscalar* interpretation — tiers are not to be thought of as distinct subregions of a single number line, but rather as distinctive number lines themselves.

Here is a moderately aggregative theory that is ecumenical between these two interpretations:

> **Fixed Moderation**: (i) There are two or more tiers such that (for every value S) all harms of severity S fall into the same tier across all possible choice situations, and (ii) distribution D_1 is more choiceworthy than D_2 just in case D_1 contains less aggregate harm at some tier and at least as little aggregate harm at every more-important tier.[17]

The idea is that we should choose whichever distribution minimizes harm at the most important tier. If two distributions do equally well at minimizing aggregate harm at the most important tier, then we choose between those two distributions so as to minimize aggregate harm at the second most important tier — and so on until the tie is broken. If the tie isn't broken, then neither distribution is more choiceworthy than the other.

It's straightforward to prove that Fixed Moderation obeys *More Ain't Less*. Let's represent different harm distributions as ordered n-tuples, $\langle \phi_1, \ldots, \phi_n \rangle$, where ϕ_i is the sum of the tier-i harm values in that distribution. (This representation is ecumenical between the mono- and multi-scalar conceptions of tiers. Given a monoscalar conception of tiers, each harm contributes a value at only one index. The multiscalar conception leaves open the possibility that a single harm might contribute a value at more than one index.) Given Fixed Moderation, if we add a number of harms of the same severity, $\langle n, \ldots, m \rangle$, to a choice situation, each of those harms will make *exactly the same contribution* to the distributions to which they belong. Each of them will change the aggregate harm value of their distribution only by *adding* $\langle n, \ldots, m \rangle$ to it. Adding *at least as many* of these harms to distribution B as to A, then, cannot possibly result in B having comparatively *less* weight along *any* dimension. This follows from the elementary

[17] This is only one way of fleshing out the Fixed view. Other choiceworthiness rules are possible. For example, we might have a Fixed rule that does not allow harms at less-important tiers to "break ties" at more-important tiers. Someone might opt for such an alternative theory if they were, unlike me, troubled by what Kamm has called the problem of "irrelevant utilities". See her often-referenced Sore Throat case (1998: 101).

mathematical fact that: for any positive integer $n \geq 1$, any $x > y$, and any $z \geq 0$, $x + n(z) > y + z$.

More Ain't Less is an extremely plausible moral principle. The fact that Fixed is consistent with this principle gives us a strong reason to favor Fixed over Variable.

5 Reasons to Prefer Variable over Fixed?

We've seen one very compelling reason to favor Fixed over Variable. Of course it might be that we should still prefer Variable *all things considered*. We should do so if there are reasons to prefer Variable that are stronger than the above reason to prefer Fixed.

In this section I consider a number of candidate reasons for favoring Variable over Fixed. I argue that none of these candidates in fact give us any reason to favor Variable over Fixed.

5.1 The Problem of Sharp Cut-offs

In §1 I noted that even amongst those committed to moderate aggregation, the Fixed view has had few proponents. I speculated as to why. The reason, I suggested, was that theorists take Fixed to require implausibly arbitrary, sharp cut-offs. They take Fixed to imply that

> Sharp Cut-Offs. For some harms x and y of only slightly different severity, you should avert any (finite) number of x harms over averting any (finite) number of x harms.

It's true that Fixed has this implication.[18] One might be tempted to think that this gives us some reason to reject Fixed in favor of Variable. But it doesn't. Variable has exactly the same result; it too implies *Sharp Cut-Offs*. Consider, for example, the Simple Rule version of Variable. On this theory, in any choice situation there is some value n such that (i) harms that are at least n-less weighty than the weightiest harm in the situation are irrelevant, and (ii) harms that are *not* at least n-less weighty than the weightiest harm in the situation are relevant.

[18] At least Fixed has this implication if we take the borders between tiers to be determinate rather than "fuzzy".

Given (i) and (ii) we get a "sharp cut-off" at the following location along the harm-severity spectrum: the weight of the weightiest harm minus n. Suppose you must choose between rescuing group A or group B. Group A consists of one person facing a harm of severity s and one person facing a lesser harm only *slightly* more weighty than $s - n$. Group B consists of one person facing severe harm and ten-million people each facing a harm of weight $s - n$.

Intuitively, you should rescue group B, since you should rescue each of ten-million people from some harm over rescuing only one person from an only *slightly* more serious harm. But because harms of $s - n$ are irrelevant in this situation, whereas harms only slightly more weighty than $s - n$ are relevant, the theory has the result that you should avert any finite number of $s - n^+$ harms over averting any number of $s - n$ harms. And thus the theory implies *Sharp Cut-Offs*.

The problem of sharp cut-offs is a problem for both Fixed and Variable. But I think it may be *less* of a problem for (certain versions of) Fixed. Given a *multiscalar* interpretation of the theory, harms of one severity lexically dominate harms of another severity only if the former has non-zero value along some dimension of well-being and the latter has zero value along that dimension. This means that, on the multiscalar interpretation, harms lexically dominate other harms only if the they differ in *kind*, and not just *degree*. Not so on Variable. Variable permits relations of lexical dominance between harms that differ only quantitatively. This fact seems to me to favor Fixed. It seems more plausible that relations of lexical dominance would arise between harms that differ in kind than between harms that differ merely in degree.

5.2 The Problem of Risk

A second problem facing Fixed is that every way of extending the account to cases involving *uncertainty* has counterintuitive implications. It might be thought that this gives us some reason to reject Fixed in favor of Variable. Let's get clear on this problem before considering whether it gives us any reason to prefer Variable.

There are a number of ways to extend Fixed to deal with uncertainty on the part of the agent. The most obvious way is to restate the Fixed account in terms of *expected* harm. On this approach, Fixed tells us to choose (between two options) the distribution that contains less aggregate *expected* harm at some tier and at least as little aggregate *expected* harm at every more-important tier.

This approach, however, has counterintuitive implications in cases like

> *Headache Relief.* Person A has a mild headache. You can either (i) allow A to suffer the headache for another hour or (ii) relieve her headache by imposing a one-in-a-zillion chance of death on B.

Intuitively, it is at least permissible to relieve A's headache, imposing a minuscule risk of death on B. We impose such risks all the time, and we think it reasonable to do so. We might, for instance, drive to the pharmacy with one of our children in order to pick up Tylenol for our other child. But consider what the expected-harm extension of Fixed must say about a case like this. In order to get the moderate verdict in cases where we can save one person from death or a great many people from mild headaches, the proponent of Fixed must say that the harm of death has weight at some tier that is more important than any of the tiers in which mild headaches have weight. If you want to minimize the *expected* amount of harm at this more important tier, then, you should choose the option that minimizes the risk of death, *no matter how small that risk*. Thus, if you want to minimize the expected value of the most important sorts of harms in Headache Relief, you should let A suffer her headache.

We might try to avoid this counterintuitive result by stipulating a *probabilistic threshold* such that we "discount" or "ignore" harms below that threshold; harms that are extremely unlikely do not figure into the aggregate value of expected harm of an option. This approach would allow us to avoid the counterintuitive result in Headache Relief, since it would have us discount the very tiny possibility of death imposed in that case.

But this move introduces different sorts of problems. Let p be the probabilistic threshold such that we ignore harms with probability less than p but take into consideration harms with probability greater than or equal to p. And now consider the following cases:

> *Relief 1*: The Xs are resting with mild headaches. You can either (i) allow each of the Xs to suffer their headaches for another hour or (ii) relieve all their headaches by imposing a $p/2$ chance of death on A.

> *Relief 2*: The Ys are resting with mild headaches. You can either (i) allow each of the Ys to suffer their headaches for another hour or (ii) relieve all their headaches by imposing a $p/2$ chance of death on A.

> *Relief 3*: The Xs and Ys are resting with mild headaches. You can either (i) allow each of the Xs and Ys to suffer their headaches for another hour or (ii) relieve all their headaches by imposing a p chance of death on A.

The adoption of probabilistic thresholds gives us bizarre results in series like this. Since we are to ignore the $p/2$ risk of harm to A in both Relief 1 and Relief 2, Fixed tells us to choose option (ii) in each of those cases. But in Relief 3 the risk of death is just great enough to make it "count". Since the expected-harm version of Fixed tells us to minimize expected harm along the most important dimension, Fixed tells you to allow the Xs and Ys to suffer their headaches in Relief 3. But of course Relief 3 is a choice between two distributions of expected harm that are *identical* to the two distributions of expected harm we must choose between when we consider Relief 1 and Relief 2 together. Relief 3 is just like the conjunction of Relief 1 and Relief 2 is all morally relevant respects. And thus if we should choose (ii) in Relief 1/2, then we should also choose (ii) in Relief 3. Amending Fixed by way of a probabilistic threshold gives us the opposite result.

We might try something other than an appeal to *expected* harm. The most obvious alternative is an appeal to *counterfactual* harm. On this approach, Fixed tells us to choose (between two options) the option that *would*, if chosen, result in less aggregate harm in some tier and at least as little aggregate harm in every more-important dimension.

But this leads to different sorts of counterintuitive results. Consider a case like

> *Parfit's Mine.*[19] One-hundred miners are trapped underground, with flood waters rising. You know that all of these men are in one of two mine shafts, but your evidence is such that you're 50/50 on which.[20] You can close one of three flood-gates. The results would be as follows:

	Shaft A	Shaft B
p	0.5	0.5
Close gate 1	0 die	100 die

[19] Parfit (1988).

[20] Assume that you are a paradigm of epistemic virtue. You've arrived at your probabilistic judgments by way of careful and rational consideration of all the evidence.

	Shaft A	Shaft B
Close gate 2	100 die	0 die
Close gate 3	1 dies	1 dies

Intuitively you should close gate 3.[21] But Fixed Moderation is forced to say otherwise on the counterfactual approach. The miners are either in shaft A or B. If the miners are in shaft A, then no one *would* be harmed were you to close gate 1, and so Fixed must say that closing gate 1 is a more choiceworthy option than closing gate 3. If, instead, the miners are in shaft B, then no one *would* be harmed were you to close gate 2, and so Fixed must say that closing gate 2 is a more choiceworthy option than closing gate 3. So Fixed implausibly implies that there is some option more choiceworthy than closing gate 3.

The lesson is that the most obvious ways of extending Fixed to cases involving uncertainty yield counterintuitive implications. This is a cost for Fixed. But this cost does not constitute a *comparative* advantage for Variable. This is because Variable suffers comparable problems in cases involving uncertainty.

Fixed and Variable suffer the same sorts of problems because these problems are general ones that plague *any* moderately aggregative theory. As Horton (2020) argues, any moderately aggregative theory is subject to a dilemma with respect to three cases involving uncertainty.

Consider, first

> *Villain*: There are one zillion X persons and one zillion Y persons. A villain will either (i) inflict a headache on each of the Xs or (ii) kill ten of the Ys.

It is non-negotiable that any moderately aggregative theory should tell you to choose option (i). The whole motivation for adopting a moderate theory in the first place is to avoid the result that one should save some number of people from minor harms over saving some number of people from very severe harms. To go in for a theory that tells you to choose option (ii) in Villain is to go in for a theory with the sorts of features of Pure Aggregation that moderate theories are specifically designed to avoid.

[21] For structurally similar cases see Regan (1980: 264f) and Jackson (1991).

Next recall

> *Headache Relief.* Person A has a mild headache. You can either (i) allow A to suffer the headache for another hour or (ii) relieve her headache by imposing a one-in-a-zillion chance of death on B.

As we noted earlier, it is intuitive that you should (or at least *may*) choose option (ii). It seems permissible to allow B to suffer a *very* tiny risk of death in order to save A from a headache.

Finally consider a series of choices in which each choice is exactly like Headache Relief:

> *Villainous Lottery*: There are one zillion X persons and one zillion Y persons. A villain pairs each X person with a Y person. For each pair the villain will either (i) inflict a headache on the X person or (ii) give the Y person a ticket for a lottery with one zillion tickets, ten of which will result in the ticket-holder's death. You must choose which.

If you should choose option (ii) in Headache Relief, then of course you should choose option (ii) in each of your one-zillion choices in Villainous Lottery. But here we run into contradiction. Taken together, the choices in Villainous Lottery constitute a choice between allowing one-zillion people to suffer a headache or ten people to die — *the very same choice you face in Villain.* Moderately aggregative theories are committed to the result that you should rescue ten people from death over rescuing one-zillion people from headaches in Villain. Since Villainous Lottery is identical to Villain in all morally relevant respects, consistency demands that moderately aggregative theories tell you to choose option (i) for every choice in the Villainous Lottery series — in contradiction to our earlier conclusion that you should choose option (ii).

To resolve this inconsistency without surrendering the central goal of the moderate project (i.e., avoiding the counterintuitive features of Pure), moderates must bite one of two bullets. One option is to reject the intuitive verdict in *Headache Relief.* That is, the moderate might accept that it is impermissible to impose even the tiniest risk of severe harm in order to avert mild harms. A second option is to reject the plausible moral symmetry between cases like Villain and Villainous Lottery. That is, the moderate might reject the plausible principle that: if you should avert harms X_1 over Y_1 and you should avert harms X_2 over Y_2 and ...

you should avert harms X_n over Y_n, then you should avert $\{X_1, X_2, \ldots X_n\}$ over $\{Y_1, Y_2, \ldots Y_n\}$.

My point is not that we should reject moderately aggregative theories in light of this dilemma. I leave it as an open question whether we should prefer to grab hold of one of these two horns or to reject moderately aggregative theories entirely. My point is just that the problems that uncertainty makes for Fixed are problems that uncertainty makes for *all* moderately aggregative theories, Variable included. There is no reason to prefer Variable over Fixed to be found here.

5.3 The Problem of Rationale

A third problem facing Fixed is that it may seem mysterious, or lacking any rationale. We would have some reason to prefer Variable over Fixed if Variable admitted of a more plausible rationale. Does it?

Let's start by looking at the rationales that have been proposed for Variable. There are two. One appeals to *respect*, the other appeals to *duties of sacrifice*. We'll consider these two rationales in turn.

The Respect Rationale. On the Variable approach, some harms "count" in a given choice situation and some harms don't. The harms that count — the "relevant" harms — are those that are close in weight to some "anchoring" harm. What explains this? Why does this closeness-in-weight relation make such an important moral difference?

According to the Respect Rationale, the closeness-in-weight relation makes an important moral difference because it makes a difference to facts about respect. We treat a person X with *disrespect* if, instead of rescuing her, we rescue any number of Ys from harms that are each significantly less weighty than the harm facing X. But we do not treat X with disrespect if, instead of rescuing her, we rescue many Ys from harms that are *not* significantly less weighty than the harm facing X. This, we are told, explains why only harms that are close in weight to the anchoring harms in a situation are relevant. Respect explains relevance.[22]

The explanation appears initially attractive to many non-consequentialists because it builds on a very popular idea: the idea that respect is central to morality,

[22]To my knowledge, this explanation has not been developed by anyone. But it is hinted at by Kamm (2007:298) and Lazar (2018:128).

and that actions should be evaluated, not just in light of how well they promote well-being, but also in light of the respect they show to others. If respect is central to morality, then it is natural to think that respect would play some role in moral aggregation.

The Respect Rationale claims a good deal more than this, however. It makes specific claims about the conditions under which respect is, or is not, shown — claims that are quite implausible on inspection.

Consider, for example, what the Respect Rationale would have to say when applied to, for example, the Simple Rule version of Variable. The Simple Rule, recall, delivers (very) surprising results in pairs of cases like Fire 1 and Fire 2. It tells us to rescue 1,000,000 people from mild harm over rescuing 20 people from moderate harm in Fire 1, but to rescue 20 people from moderate harm over rescuing 1,000,000 people from mild harm *and* 1 person from severe harm in Fire 2.

	Building A		Building B	
Fire 1	moderate	x20	**mild**	**x1,000,000**
Fire 2	**moderate**	**x20**	mild	x1,000,000
			+ severe	x1

If *respect* is to explain this result, one of two things would need to be true. Either

(i) You would *not* disrespect any of the 20 people facing moderate harm by saving 1,000,000 people from mild harm instead of them in Fire 1. But you *would* disrespect each of the 20 people facing moderate harm by saving 1,000,000 people facing mild harm plus 1 person facing severe harm instead of them in Fire 2.

or

(ii) You would disrespect each of the 1,000,000 persons facing mild harm by saving 20 people from moderate harm instead of them in Fire 1. But you would *not* disrespect each of the 1,000,000 people facing mild harm by rescuing 20 people from moderate harm instead of them in Fire 2.

Either option is extremely implausible. If you disrespect someone in virtue of saving a certain group of people instead of her, then surely you disrespect her *at*

least as much when you save a *subset* of that group instead of her.

To be clear, I'm not denying that respect facts are central to morality, nor am I deny thing that facts about respect and facts about aggregation might be closely connected.[23] It's very plausible that they are. What's not plausible are the specific claims about respect that must be made if one wants facts about respect to explain Variable.

The Duty Rationale. Voorhoeve (2014) proposes an alternative — and I think, more interesting — way of explaining why only harms that are close in weight to some anchoring harm(s) are "relevant". His explanation begins with the widely-accepted thought that morality permits a *limited* amount of self-preference:

> Commonsense morality judges it permissible to save yourself rather than a stranger from an *equally large* harm. It also regards it as permissible to save yourself from a *lesser* harm rather than a stranger from a somewhat greater harm. For example, if you can either save yourself from [...] complete disablement or save a stranger from death, then you are permitted to prevent your disablement. But there are limits to the extra concern for your dear self that everyday morality permits. If you face a very minor harm such as an illness that will leave you bedridden for a day and can either prevent this harm to yourself or prevent the death of a stranger, then it holds that you must save the stranger (71).

Morality, says Voorhoeve, permits a person to privilege her own rescue over a stranger's just in case the harm she is threatened with is not significantly outweighed by the harm threatening the stranger. It is here that the *closeness-in-weight* relation enters the story.

> **V1**. (Ceteris peribus) A is duty-bound to suffer harm h in order to rescue a stranger from harm just in case h is significantly less weighty than the harm facing the stranger.

The next part of Vorhooeve's explanation connects these duties of sacrifice to a person's "standing to complain". Suppose A is in danger of suffering harm h_A and B is in danger of suffering harm h_B. And suppose you can avert one harm or the other. In such a situation, says Voorhoeve, whether A would have a

[23] Of course, it might be that facts about respect are explained in terms of facts about aggregation, rather than the other way around.

complaint if you rescued B depends on whether A would have a duty to bear cost h_A in order to rescue B from h_B, were A able to do so.

> **V2**. (Ceteris peribus) If A *would* have a duty to bear a harm h_A in order to rescue a stranger B from a harm h_B *were* A were able to do so, then A cannot complain if some stranger C rescues B from h_B over rescuing A from h_A (when C cannot rescue both A and B).

From **V1** and **V2**, we get

> **V3**. If the severity of the harm facing B is *significantly greater than* the severity of the harm facing A, then A cannot complain if some stranger C rescues B instead of A (when C cannot rescue both A and B).

If we think that the choiceworthiness of a distribution is determined, not by aggregating harms themselves, but by aggregating complaints, then we arrive at an explanation for why only harms that are close in weight to some anchoring harm(s) are "relevant".

Like with the respect rationale, the Vorhooeve explanation begins with an appeal to a familiar and compelling moral idea. Most people find it plausible that morality allows some amount of personal preference. And most people find it plausible that morality does not allow *unlimited* personal preference. Some harms to strangers are serious enough to put me under a duty of rescue that may require me to bear lesser costs.

I don't question this familiar idea. But I think there is good reason to be dubious about Vorhooeve's *application* of this idea.

I find V2 especially untenable. V2 claims that the following inference is a good one:

> A *would* have a duty to bear a harm h_A in order to rescue a stranger B from a harm h_B were A *were* able to do so.

Therefore,

> A cannot complain if some stranger C rescues B from h_B over rescuing A from h_A when C cannot rescue both A and B.

Vorhooeve here posits a rigid symmetry between first-personal and third-personal

considerations. He takes the view that when an agent is deciding which persons to save from harm, she should take their interests into consideration in *the same way* that those strangers should take their own interests into consideration.

I find such a rigid symmetry implausible. Consider the following pair of cases:

> *On Duty.* When A joined the Coast Guard she took an oath to rescue anyone who needed rescuing while she was out on patrol, even at somewhat more serious costs to herself. While A is on patrol one day, an unexpected storm strikes, seriously damaging her boat, and rendering her unconscious. She will suffer a broken arm and a broken finger if not rescued. Nearby, a second boat has also been damaged, and its passenger, B, will suffer a broken arm if not rescued. C is in the area and can rescue either A or B, but not both.
>
> *Off Duty.* As before, except that A has just gotten "off work" when the storm strikes.

In On Duty, if A were able to rescue B, she would have a promissory and occupational duty to do so, even if it meant suffering both a broken arm and broken finger. Not so in Off Duty. Since A is "off duty" in that case, she is under no promissory or occupational duty to suffer a greater harm in order to rescue B from a lesser harm.

The Vorhooeve explanation thus delivers the following implausible result. It tells us that C should rescue A rather than B in Off Duty, but that C should effectively *ignore* A's interests in On Duty, and rescue B from the lesser harm. I find this implausible on two counts. First, it seems counterintuitive that A's interests should be ignored in On Duty on account of the duties she *would* have were she conscious and capable of rescuing B. But even more seriously, I find it wildly implausible that C's moral reasoning should be so significantly constrained by whether A's shift at work had just ended or not at the time that the storm struck. It seems bizarre that A's interests could carry the same weight as B's interests the moment after A's shift ends, but that A's interests could be lexically inferior to B's interests the moment before her shift ends.

It isn't just "acquired" duties (such as promissory duties) that break the rigid symmetry required by V2. Natural duties also break the mold. Consider the following pair of cases:

> *Strangers Fishing.* A (a woman of forty years old) is out fishing at sea. Nearby, B (a boy of twelve years old) is also fishing. An unexpected storm strikes, rendering them both unconsious. If no one helps, B will suffer a broken arm, and A will suffer a broken arm *and* a broken finger. C, a stranger, is the only person who can provide aid. But she only has the resources to save either A or B from their potential injuries.
>
> *Family Fishing.* As before, except that A is B's mother.

Parental duties are quite demanding. Plausibly, if A were able to rescue her son from a broken arm at the cost of a broken arm and broken finger, she would have a duty to do so (as in Family Fishing). But of course our duties towards strangers are not so demanding. A would not have a duty to bear a greater cost in order to rescue a *stranger* from a lesser cost (as in Strangers Fishing).

So again the Vorhooeve explanation delivers up an implausible result. It tells us that C should weigh A's interests against B's in Stranger Fishing, but that C should effectively *ignore* A's interests in Family Fishing. Again, this seems implausible on two counts. First, it is counterintuitive that C should effectively ignore A's interests in Family Fishing. A's special duty to prioritize her son's interests is the product of *her* special relationship to her son. C does not stand in such a relationship to B. So why should C's moral reasoning be bound by the same factors that bind A's? Second, it is implausible that the extent to which C should consider A's interests should vary *so much* between the two cases. Perhaps the fact that A is B's mother should make *some* difference. But why would it make *all* the difference, rendering A's interests irrelevant?

V2 is implausible. It implies much too rigid a symmetry between first-personal and third-personal considerations.[24]

Neither the appeal to respect, nor the appeal to counterfactual duty, offers us any

[24] I think there is room to push on V1 as well. V1 is intended to capture the intuitive idea that morality allows for limited personal preference. But V1 is only one precisification of this rough idea. V1 claims that we should think about the limits of personal preference in terms of differences in *degree*. We might instead, however, precisify the idea in terms of differences in *kind*. It might be that I incur duties of sacrifice, not when the cost of my sacrifice is sufficiently lesser in degree than the harm facing some stranger, but when the cost of my sacrifice is sufficiently lesser in kind. On this alternative precisification, the Voorhoeve-style appeal to duties of sacrifice would deliver us Fixed rather than Variable. The upshot: absent an independent argument for precisifying the idea of agent-relative prerogatives in terms of differences in degree rather than differences in kind, this idea does not support Variable over Fixed — even if morality conformed to the rigid symmetry demanded by V2.

reason to prefer Variable over Fixed. Now perhaps we will eventually discover a compelling rationale that applies to Variable (but not to Fixed). But I see no reason we should expect such an explanation to be more forthcoming for Variable than for Fixed. If anything, there are reasons to think the project of explaining Fixed *more* promising. Fixed has a much simpler structure, which admits simpler explanations. Fixed only requires us to explain why there are different *kinds* of well-being such that some kinds take lexical, deontic priority over others.

By way of example, one straightforward way we might explain why some kinds of well-being take lexical, deontic priority over others is by appeal to a lexically-ordered *axiology*. It might be that we *should* rescue one person from death over rescuing any number of people from a mild headache simply because it is *better* to avert one death than to avert any number of mild headaches.[25] This is not an outlandish idea. As Parfit (1984:388) notes, we intuitively judge it better that (say) one-billion people live excellent lives than that any number of people live lives that are barely worth living. And as McTaggart (1927:452-453) notes (following Mill), we intuitively judge it better that one person live a flourishing life of "knowledge, virtue, love, pleasure, and intensity of consciousness" for a million years than that one "oyster-like" creature of little consciousness live a painless life of very mild pleasures for any length of time. There are different ways we might accommodate these intuitions, but one of the best explanations is that well-being has a lexical structure.

Of course this only pushes the explanatory question back a level. We should now want to explain this axiology; we should now want to know why some amounts of some kinds of well-being are *better than* any amounts of certain other kinds of well-being. But it isn't unreasonable to think we might be able to explain this as well. There are well-known theories of well-being that are conducive to such an idea. For example, many theories attempt to explain well-being in terms of *desire*. "Actual desire" accounts attempt to explain a person's well-being in terms of her actual desires; "natural desire" accounts attempt to explain a person's well-being in terms of desires that are characteristic of her species; "ideal desire" accounts attempt to explain a person's well-being in terms of the

[25] Aristotle, Mill, Brentano, and Ross each arguably endorsed a lexically-ordered conception of welfare or axiology (see Lemos 1993). For more contemporary lexical conceptions of welfare or axiology, see Kitcher (2000), Nebel (2021), Nussbaum (1995), Portmore (2000), Sen (1980), and Thomas (2017).

desires she *would* have under certain ideal conditions; "ideal observer" accounts attempt to explain a person's well-being in terms of what an idea observer would desire for her. Any of these theories would deliver lexical dominance between different kinds of well-being if real or ideally-situated people had lexically-ordered desires — i.e., if they would prefer some amount of gain with respect to certain capacities, pleasures, or experiences over any amount of gain with respect to other capacities, pleasures, or experiences.

But I digress. The point is not that we can or should explain the Fixed view in terms of desires. The point is just that there are possible avenues by which Fixed *might* be explained, and that the prospects for explaining the view seem (as of now) at least as promising as the prospects for explaining Variable.

6 Conclusion

Many theorists want to be moderate about aggregation. They want a theory that allows *some but not all* inequivalent harms to trade-off against one another. At present, the moderate theories discussed in the literature are all variants of the Variable view. They all appeal to a non-transitive *closeness-of-weight* relation in order to sort the harms in a given choice situation as "relevant" or "irrelevant".

In recent years we have seen increasingly sophisticated versions of this view. But, as I've argued, all of this sophistication accomplishes little. *Every* version of Variable — from the most simple to the most complex — violates the principle *More Ain't Less*. Fixed, however, obeys this compelling principle. What's more, what reasons we might have thought there were to prefer Variable over Fixed turn out, upon inspection, not to favor Variable at all. There are compelling reasons to prefer Fixed; there are no compelling reasons to prefer Variable; and thus we should prefer Fixed over Variable.

List of Works Cited

Anscombe, Gertrude E.M. "Who is Wronged?" Oxford Review 5 (1967): 16-17.

Doggett, Tyler. "Saving the Few." Noûs 47 (2013): 302–315.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and

Dearest Objection." Ethics 101 (1991): 461-82.

Halstead, John. "The Numbers Always Count." Ethics 126 (2016): 789-802.

Horton, Joe. "Aggregation, Complaints, and Risk." Philosophy & Public Affairs 45 (2017): 54-81.

— "Always Aggregate." Philosophy & Public Affairs 45 (2018): 160-174.

— "Aggregation, Risk, and Reductio." Ethics 130 (2020): 514-529.

Kamm, Frances. Morality, Mortality Volume I. Oxford University Press, 1998.

— "Aggregation and Two Moral Methods." Utilitas 17 (2005): 1–23.

— Intricate Ethics. Oxford University Press, 2007.

Kitcher, Philip. "Parfit's Puzzle." Nous 34 (2000): 550-577.

Lazar, Seth. "Limited Aggregation and Risk." Philosophy & Public Affairs 46 (2018): 117-159.

Lemos, Noah M. "Higher Goods and the Myth of Tithonus." Journal of Philosophy 90 (1993): 482-496.

McTaggart, John M.E. The Nature of Existence Vol. 2. Cambridge University Press, 1927.

Munoz-Dardé, Véronique. "The Distribution of Numbers and the Comprehensiveness of Reasons." Proceedings of the Aristotelian Society 105 (2005): 191–217.

Nebel, Jacob M. Totalism without Repugnance. In Ethics and Existence: The Legacy of Derek Parfit, edited by Jeff McMahan, Tim Campbell, James Goodrich, and Ketan Ramakrishnan. Oxford University Press, 2021.

Norcross, Alastair. "Comparing Harms: Headaches and Human Lives." Philosophy & Public Affairs 28 (1997): 135-167.

— "Contractualism and Aggregation." Social Theory and Practice 28 (2002): 303-314.

— "Two Dogmas of Deontology." Social Philosophy and Policy 26 (2009): 76-95.

Nussbaum, Martha. “Human Capabilities, Female Human Persons.” In Women, Culture, and Development, edited by Martha Nussbaum and Jonathan Glover. Oxford University Press, 1995.

Otsuka, Michael. “Saving Lives, Moral Theory, and the Claims of Individuals.” Philosophy and Public Affairs 34 (2006): 109-135.

— “Risking Life and Limb.” In Identified versus Statistical Lives, edited by Glenn Cohen, Norman Daniels, and Air Eyal. Oxford University Press, 2015.

Parfit, Derek. Reasons and Persons. Oxford: Clarendon Press, 1984.

— “What We Together Do,” unpublished manuscript, 1984.

Portmore, Douglas. “Does the Total Principle Have any Repugnant Implications?” Ratio 12 (1999): 80-98.

Regan, Donald. Utilitarianism and Cooperation. New York: Oxford University Press, 1980.

Reibetanz, Sophia. “Contractualism and Aggregation.” Ethics 108 (1998): 296-311.

Scanlon, Tim. What We Owe to Each Other. Belknap Press of Harvard University Press, 1998.

Sen, Amartya. “Equality of What?” In Liberty, Equality, and Law: Selected Tanner Lectures on Moral Philosophy, edited by John Rawls and Sterling McMurrin. University of Utah Press, 1987.

Shiffrin, Seana Valentine. “Harm and Its Moral Significance.” Legal Theory 18 (2012): 357-398.

Steuwer, Bastian. “Aggregation, Balancing, and Respect for the Claims of Individuals.” Utilitas 33 (2021): 17-34.

Tadros, Victor. "Localized Restricted Aggregation.’ In Oxford Studies in Political Philosophy Volume 5, edited by David Sobel, Peter Vallentyne, and Steven Wall. Oxford University Press, 2019.

Taurek, John. “Should the Numbers Count?” Philosophy and Public Affairs 6 (1977): 293-316.

Thomas, Alan. "Giving Each Person Her Due: Taurek Cases and Non-Comparative Justice." Ethical Theory and Moral Practice 15 (2012): 661–676.

Thomas, Teruji. "Some Possibilities in Population Ethics." Mind 127 (2017): 807-832.

Tomlin, Partick. "On Limited Aggregation." Philosophy & Public Affairs 45 (2017): 232-260.

van Gils, Aart & Tomlin, Patrick (2020). "Relevance Rides Again? Aggregation and Local Relevance." In Oxford Studies in Political Philosophy Volume 6, edited by David Sobel, Peter Vallentyne & Steven Wall. Oxford University Press, 2020.

Voorhoeve, Alex. "How Should we Aggregate Competing Claims?" Ethics 125 (2014): 64-87.

— "Balancing Small Against Large Burdens." Behavioural Public Policy 2 (2018): 125-142.