The Rank-Matching Theory[*]

Abstract

This article defends a new account of how the interests of separate people bear on what to do: Rank-Matching Theory. Unlike 'partial' or 'limited' aggregation, it wholly rejects the idea that the interests of separate people sum. Unlike traditional 'anti-aggregative' theories, it is nonetheless sensitive to the numbers.

We argue that the commitments underlying this theory make better sense of our intuitive reactions than competitors, and we prove that the theory follows from some natural principles. We also use simulations to show how, given certain empirical claims, it may be able to explain the distinction between 'relevant' and 'irrelevant' interests. We conclude the Rank-Matching Theory merits further investigation.

Word count: 108 (abstract), 10,354 (body and notes), 1,396 (appendix).

One basic disagreement in moral philosophy concerns whether the boundaries between the lives of separate people have any moral significance. Certain examples make the dispute vivid:

> Suppose that Jones has suffered an accident in the transmitter room of a television station. Electrical equipment has fallen on his arm, and we cannot rescue him without turning off the transmitter for fifteen minutes. A World Cup match is in progress, watched by many people, and it will not be over for an hour. Jones's injury will not get any worse if we wait, but his hand has been mashed and he is receiving extremely painful electrical shocks. Should we rescue him now or wait until the match is over? (Scanlon 1998: 235)

Let us grant, *arguendo*, that if we took the irritation or disutility of each audience member and summed them all together, the result would exceed what Jones stands to suffer. No different,

[*] Claude made Figures 1, 2, and 3 in LaTeX and edited code in Footnote 31. All code was reviewed by hand.

perhaps, from how taking a penny from each audience member would yield a balance that exceeds Jones's net worth. Should we therefore abandon Jones? Some, like the utilitarians, answer in the affirmative. Others disagree. It somehow seems to matter how the additional suffering is distributed across the borders that separate one from another.

Recent moral philosophy has no dearth of theories for this and related cases.[1] Most of these theories share two features. First, they include an aggregative component, such as that the claims of separate people sum.[2] Second, they distinguish claims that are "relevant" to a given choice from claims that are "irrelevant." A claim is irrelevant when it is far weaker than the strongest claim with which it competes, and irrelevant claims have no bearing on what to do.[3]

Irrelevance explains why we must save Jones: his plight renders the irritation of each audience member irrelevant, and after removing them from consideration, there is no reason not to save Jones (and strong reason to do so). Aggregation explains why we must nonetheless save two lives rather than one. On such views, the borders between people play a role at the initial, qualifying stage: they determine whether some claim gets to be considered at all. But once we have settled which claims should be considered and which should be ignored, then the borders go away and we pool together what remains on either side.[4]

Our view begins from a different diagnosis of Jones's predicament. Rather than taking his suffering to render the irritation of the many irrelevant, we claim that the irritation of the many simply does not matter very much. One can see this by considering a variant of the case that removes Jones entirely. Suppose that the transmission is delayed simply due to a solar flare. How morally important does the delay now seem? Our sense is that the delay continues to matter little. It just does not matter much, from the moral point of view, that billions of people had to wait about fifteen minutes for the World Cup transmission to resume. It is a cause that is hard to get

[1] See, e.g., Dorsey 2009, Henning 2023, Sobjectivity 2, 3, and 4 in Kamm 1993: 165–186, Liao and Lim 2024, Scanlon 1998: 229–241, Scanlon 2021, Steuwer 2021, Temkin 2005, Voorhoeve 2014, and Zhang 2024.

[2] Voorhoeve 2014. Others sum interests (Rüger 2020), votes (Henning 2023), or impersonal goodness (Zhang 2024).

[3] This is Voorhoeve 2014's initial formulation, which Tadros 2019 and van Gils and Tomlin 2020 further refine.

[4] This is so, at least, when all claims are relevant to one another.

worked up about. But it does matter if one person suffers an hour of extremely painful electrical shocks. That is why, on our view, we must save Jones. We posit no obstacle to considering the irritation of the many; rather, such irritation just does not amount to very much.

This alternative diagnosis fits a larger worldview where the interests of separate people simply fail to sum, so that there is no sense in which the many irritations together outweigh the suffering that Jones may face. This possibility takes seriously the borders between the lives of separate people; indeed, makes them permanent fixtures of the moral landscape. One aim of the paper is to place this non-aggregative worldview, which owes much to Taurek 2021, clearly on the table.

Taurek drew a notorious conclusion from such commitments: it is morally permissible to save one life rather than five.[5] We take a different route through the same terrain. Although the non-aggregative worldview significantly restricts the available theoretical resources, thereby constraining the space of possible views, it does not uniquely imply Taurek's view, and the paper proposes an alternative. We start with the general idea, shared by many contractualists, that one can justify failing to help someone by appealing to the competing interests of someone else.[6] We develop this idea in a new way.

When choosing between two options, we ask the person with the most to gain from one option to consider the perspective of the person with the most to gain from the alternative, and we ask the person with the *second-most* to gain from one option to consider the person with the *second-most* to gain from the alternative, and so on. We match people by how their interests rank, with stronger interests matched with stronger interests and weaker with weaker.

The perspective of each then provides some justification to their counterpart for not choosing the latter's favored action. When one's interests are stronger than those of one's counterpart, one is left with a "residual complaint" whose strength corresponds to the difference between one's interests and those of one's counterpart. We ought to minimize the maximum residual complaint.

[5]Taurek 1977: 309–310.

[6]See, e.g., Scanlon 1982: 113.

Call this the Rank-Matching Theory. It subtracts, but does not sum, individuals' interests.

Our primary case for the Rank-Matching Theory goes via the preceding path: Jones's predicament raises doubts about aggregation. Combining these doubts with a prohibition on double-counting and an obligation to minimize the maximum complaint implies the theory.

Our secondary case appeals to the implications of the Rank-Matching Theory. We show how, given certain normative and empirical assumptions, the Rank-Matching Theory makes plausible prescriptions and promises to explain the relevance and irrelevance relations that other views take as primitive.[7] We begin by surveying the implications of the Rank-Matching Theory when applied to a single choice considered in isolation from any other choice one has faced or will face. In this isolated setting, Rank-Matching Theory has a mixed record that includes both plausible and implausible consequences that its rivals lack.

But isolation is not realistic. Often, an agent makes many choices that affect the same people over time. It then seems reasonable, not to consider an individual choice in isolation, but to consider the whole series of decisions that one faces. We present a simple model of the choices that an agent faces over time, and we show how the Rank-Matching Theory's prescriptions in this model avoid the implausible consequences that arise in the isolated setting. It remains true, though, that one must secure the far stronger interests of the one rather than the far weaker interests of the many. A limited form of aggregation thus emerges simply from adding other choices to the model. We flag, though, that our model makes significant idealizations for tractability, such as setting aside uncertainty about what one's choices will bring about.

Here is the structure of the paper. We begin with our diagnosis of the Jones case and our non-aggregative commitments (§1). Next, we present the Rank-Matching Theory as a natural development of contractualist and non-aggregative ideas (§2). We consider its implications for a single choice considered in isolation (§3) and then its implications in a simple model of a series of choices (§4). We close with a final prediction of the theory.

[7] Tadros 2019: 178. Zhang 2024: 490 - 492.

1 Foundations

We assume that one ought to save Jones in the transmitter room.[8] This suggests a more general principle that we also take for granted:

> Disparate Interests
>
> If someone has interests that are far stronger than anyone's competing interests, then all else equal, one ought to help them.

where "far stronger" should be understood along the lines of the enormous disparity between Jones's interest in avoiding an hour of extremely painful electrical shocks and a viewer's irritation at a fifteen minute delay. As mentioned, there are two explanations of why we must save Jones:

- Irrelevance Explanation: Jones's plight renders the irritation of the audience members irrelevant, so they have no bearing on what to do. But if their irritation was relevant, then collectively their interests would matter far more than those of Jones.
- Unimportance Explanation: the irritation of the audience members does not matter much, whereas Jones's plight does. So we ought to help Jones.

Both explanations make the same prediction in Transmitter Room, but they disagree in variations of the case and we can use these variations to adjudicate among them. Consider:

> Overtime
>
> You are an engineer at the television station. Having just gotten home after an exhausting eighteen hour day at the office, you realize that a bug in the code will cause a fifteen-minute delay in tomorrow's World Cup transmission, irritating each of the many viewers. The bug depends on subtle dependency interactions; it is the kind of thing no one could be expected to find, and that no one will ever know that you knew. To prevent it, you would have to take the subway thirty minutes back to the office and fix the bug yourself.

and

[8] For an overview of the debate over this judgment, see Horton 2021.

Overtime (Shocks)

> As before, except that the bug sharply increases the voltage in the transmitter room. As a result, the first employee to arrive, Jones, will suffer extremely painful electrical shocks for an hour until others trickle into the office and find him. Again, the bug is extremely subtle and no one will ever know that you knew. To prevent it, you would have to take the subway thirty minutes back to the office and fix the bug yourself.

Here are our reactions to these cases. In Overtime, perhaps you really should go back to the office and fix the bug. But it also seems completely understandable if, exhausted, you just want to collapse for the night. Tomorrow, no doubt, the sports section will run a short story about the unexpected delay, and network executives will be furious. In a few years, though, you will regale friends and anyone who will listen at the bar about how you saw the delay coming and just couldn't be bothered to fix it. They will find your story fascinating, or at least pretend to.

In Overtime (Shocks), you must go back to the office. To take a line from Scanlon, not doing so "would be inhumane even if it were not actually wrong" (Scanlon 1998: 238). If you stay at home, tomorrow the paper will run a harrowing piece about the experience of poor Jones. You will never admit your part in this story to your closest friends, let alone strangers.

These reactions conflict with what the Irrelevance Explanation predicts. In Overtime, presumably the irritation of the audience members is relevant to your decision to go back to the office or not. (Here, the disparity between your interests and theirs is much smaller than in the case of Jones.) So we should aggregate their interests when deciding what to do, whereupon their interests far outweigh yours and you must go back to the office. Given what they together have at stake, not going back is seriously wrong. But this is not how the case strikes us intuitively.

Alternatively, if the irritation of an audience member is not relevant to your decision, then the Irrelevance Explanation predicts that you ought to stay at home and that nothing even counts in favor of going to the office. Again, this does not seem true. There is a reason to go to the office: billions of people will be irritated by the delay. Perhaps this reason even carries the day and you really should go into the office.

The problem with the Irrelevance Explanation is that its prescriptions are implausibly Manichean. Either the irritation of an audience member is relevant to your choice or it is not. If it is relevant, then you have incredibly strong reason to prevent the delay, and this reason is far stronger than your reason to prevent Jones's suffering. Moreover, failing to comply with such a powerful reason is presumably seriously wrong. But if it is not relevant, then you have no reason whatsoever to go to the office. The irritation of the audience members simply does not bear on the case. Neither possibility matches our intuitive sense that perhaps you should go back, but it is not seriously wrong if you do not.

The Unimportance Explanation predicts our intuitive reactions to these cases. In Overtime, the delay is just not that important. But neither is the cost, to you, of staying up for another hour and a half to fix the bug. So perhaps you really should go back in, but the stakes are not enormous either way. This accounts for our relative equanimity towards your decision. But an hour of extremely painful electrical shocks for Jones does matter, which is why you must return to the office in Overtime (Shocks). Our reactions favor Unimportance Explanation over Irrelevance Explanation.

At this point, one may worry that the Unimportance Explanation goes too far. Surely it would be reasonable to reprimand or fire someone who delayed the transmission merely out of carelessness, and it would be reasonable to take great precautions to ensure that such a popular broadcast goes off without a hitch. But if delays to the broadcast are so unimportant, how could these measures be reasonable?

Our answer is that many of the competing considerations are also unimportant when compared to things like the possibility of someone suffering an hour of extremely painful electrical shocks. It is worth remembering that most of our days are organized around things that are unimportant in this sense: watching a good movie, seeing some friends, pursuing a promotion. All of these pale in comparison to the extraordinary stakes of Scanlon's case. But ordinarily, we face tradeoffs among relatively unimportant considerations, and it can be reasonable to take great precautions for the sake of the unimportant when the costs are paid in terms of the unimportant. We return to this point in §4.3.

1.1 The Allure of Non-Aggregation

Our diagnosis of Transmitter Room already suggests that the interests of separate people do not sum. For if they did, then the combined significance of the irritation of billions of people would presumably matter enormously, pace the Unimportance Explanation. We add here a second reason for doubting that interests sum. We offer it less as knock-down argument than as a path into the non-aggregative worldview.[9] We aim to help in seeing clearly from this point of view.

Begin with some claims that we take for granted. We can compare the lives of individuals and observe that one person led a more fulfilling, exciting, pleasurable, virtuous, consequential, etc. life than another. We can also compare people in their lifetime wellbeing, or how well off they are over the course of their whole lives. Hume had a better life than Robespierre, for example. And we can make more localized comparisons, as we did above: Jones's interests in avoiding extreme pain far exceed a viewer's interests in avoiding a fifteen-minute delay. We can consider what someone has to gain or lose in some choice and compare it to the stakes for someone else. We seem to reach these conclusions by taking up the perspective of one or another and comparing them. We use empathy, at least in one sense of that overfreighted word. By taking up one perspective and then another, we learn how they differ.

Nothing similar is available when we consider the sum of the interests of others. We can imagine the irritation of one audience member at the delay. We can imagine the irritation of another audience member. We can compare these, finding perhaps that the former is more irate than the latter. But what would it mean to combine them? How could we put the irritation of one audience member together with the irritation of another? Empathy allows us to make com-

[9] For related doubts about whether the interests of separate people sum, see Korsgaard 2013 and Nebel 2023.

parisons among the interests of separate individuals, but there seems to be no mental operation that stands to aggregation as empathy stands to these comparisons.

In response to such concerns, philosophers and economists have suggested different candidates. Perhaps, for example, we should consider the perspective of one person who initially experiences the former irritation and then experiences the irritation of the latter. Or, perhaps, we should consider the position of someone who has some chance of undergoing one or another irritation and consider what lottery would be in their best interests. It would take us too far afield to consider such candidates in detail.[10] Rather, what we want to convey is the clear sense that something very different is happening here. Even if it turns out that we ought to guide our behavior by these theoretical constructions, so that when deliberating we ought to compare Jones's agony to, say, the tedium of waiting fifteen billion minutes for the World Cup transmission to resume, the point remains that they are constructions. What is real, one wants to say, are just the facts about each individual and what they stand to gain or lose.[11]

Here the metaphor of weight, commonly used for the strength of our interests or reasons, threatens to mislead. There is a clear sense in which we can sum the weights of distinct objects, and this may tend to make us think that there is a clear sense in which we can sum the interests of separate people, placing them all together on some cosmic balance beam. But this metaphor, as Nebel 2023 emphasizes, must be argued for rather than assumed. Other metaphors are no less appealing, and Taurek 2021 suggests the metaphor of beauty.

To a first approximation, the non-aggregative view of the interests of others can be taken up by thinking of others' interests as akin to the beauty of different paintings. There is a rich body of evaluative thought and talk about the beauty of paintings. We can discuss what features of some work make it so striking, say. And we can draw comparisons; one painting may be more grotesque, unsettling, thrilling, boring, etc. than another. One painting may be far more beautiful

[10]Nebel 2023 rejects several proposals like the former, while Weymark 1991 criticizes a version of the latter.
[11]See Taurek 2021: 314 for a forceful statement of this point.

than another. Yet there is no temptation to think there is a sum of the beauty of separate paintings.

1.2 The Constraints of Non-Aggregation

We propose to be guided by the analogy to beauty in thinking about the contours of a non-aggregative moral theory. But what exactly does this analogy suggest? The issue is finicky but critical. Stringent constructions can wrongly deprive non-aggregative theories of pertinent moral distinctions; lax constructions can wind up admitting aggregation through the back door. Our characterization proceeds in two parts: positive and negative.

Positively, in the case of beauty, we have:

- Beauty Level: Painting P is more beautiful than Q
  (symbolically: $P \succ Q$)

- Beauty Difference: Painting P exceeds Q in beauty by more than R exceeds S in beauty
  (symbolically: $P - Q \succ R - S$)

Beauty Difference may be initially unfamiliar, but claims like "P is far more beautiful than Q" and "R is only a little more beautiful than S" can be true, and such claims seem to imply that P exceeds Q in beauty by more than R exceeds S. Analogizing to interests:

- Interpersonal Interests: Alice's interests in having x obtain rather than y exceed Bob's interests in having z obtain rather than w
  (symbolically: $x_{\text{Alice}} - y_{\text{Alice}} \succ z_{\text{Bob}} - w_{\text{Bob}}$)

- Interpersonal Difference of Interests: Alice's interests in having x rather than y exceed Bob's interests in having z rather than w by more than Carol's interests in having $\hat{x}$ rather than $\hat{y}$ exceed David's interests in having $\hat{z}$ rather than $\hat{w}$
  (symbolically: $(x_{\text{Alice}} - y_{\text{Alice}}) - (z_{\text{Bob}} - w_{\text{Bob}}) \succ (\hat{x}_{\text{Carol}} - \hat{y}_{\text{Carol}}) - (\hat{z}_{\text{David}} - \hat{w}_{\text{David}})$)

As before, Interpersonal Difference of Interests may be less familiar, but claims like "Alice has far more at stake than Bob" and "Carol has a bit more at stake than David" make sense, and they seem to imply that Alice's interests exceed Bob's interests by more than Carol's interests exceed David's interests.

We highlight such comparisons because our theory later relies on them. Comparisons like Interpersonal Interests, for example, would not suffice. This is one place where it is imperative to clarify the full sweep of resources available to non-aggregative moral theories. Restricting such theories to comparisons like Interpersonal Interests would yield a blinkered range of possibilities.

Negatively, in the case of beauty we have:

- No Sum of Beauty: There is no sum of the beauty of separate paintings.

which we analogize to:

- No Sum of Interests: There is no sum of the interests of separate persons.

Barring explicit sums, though, is not sufficient. Because addition and subtraction are so closely related, aggregative views can be reformulated to avoid explicitly mentioning any sum. Sums can be replaced by differences. Schematically, the problem is that an aggregative theory might consider whether:

$a > b + c + d$

a exceeds the sum of b, c, and d

or it might consider whether:

$(a - b) - c > d$

The difference between (the difference between a and b) and c exceeds d.

Prohibiting only explicit summation will prohibit the former while allowing the latter, but it is not clear how much they differ. For example, I might want to know whether my atlas weighs more than my backpack, cat, and dog. Instead of directly weighing the former and the latter and then comparing, I could first determine by how much my atlas outweighs my backpack, then subtract the weight of my cat, and finally compare the result to the weight of my dog. Even if we only ever used the latter procedure, weight would still be an aggregative quantity.

These are thorny issues, especially given well-known mathematical results that use comparisons of differences to establish aggregative structures.[12] We take the example of beauty to show that comparisons of differences are compatible with the rejection of aggregation. It seems undeniable that one painting can be far more beautiful than another or only a little more beautiful, and that the former involves a greater difference in beauty than the latter. But we take this last worry to show that differences, when sufficiently rich and well-behaved, differ little from aggregation. It seems Pyrrhic to insist that there is no sense in which a million doodles are together more beautiful than Turner's *Rain, Steam, and Speed* while accepting that *subtracting* the beauty of a million doodles from that of Turner's masterpiece would leave the former with none.

This paper will not resolve the question of when exactly differences are sufficiently rich and well-behaved to amount to aggregation. But this concern is one reason why our view later considers comparisons between differences of interests (as in Interpersonal Difference of Interests) but does not consider comparisons between differences of differences of interests (or any further iterations).[13] This is a place where an overly generous construal of the resources available to a non-aggregative moral theory would void the distinction of its interest.

To summarize, we take the analogy to beauty to license comparisons like: Alice's interests in an appendectomy exceed Bob's interests in an aspirin by more than Carol's interests in going to

[12] The proof of Theorem 1 in Chapter 4 of Krantz et al. 1971: 148 - 149 shows how certain assumptions about comparisons of differences between members of some set imply that we can represent the members of that set with an additive function. See also Suppes and Winet 1955.

[13] The other reason is that we do not seem committed to such further iterations by ordinary moral discourse.

the library exceed David's interests in going to the park. And we take this analogy to prohibit not only explicitly aggregative considerations, such as the sum of Alice and Bob's interests, but also systems of differences sufficiently rich and well-behaved so as to amount to aggregation.

2 Theory

Non-aggregation constrains what ingredients our theory can use, but there are still many ways to combine them. For example, perhaps we ought to maximize the minimum level of wellbeing. Or perhaps we ought to satisfy the strongest individual interest. These possibilities and more comport with the strictures of non-aggregation.

The core of our view is contractualist and belongs to the same family as Scanlon 1982: we justify an action to an individual by asking them to consider the competing interests of another individual. Two features distinguish our theory. First, a justification can weaken the force of someone's complaint even when the justification does not fully extinguish it.[14] Second, the same person's interests cannot weaken or quell the complaints of multiple others. This is the key feature that makes our theory sensitive to the numbers, and it will be important for us to show that our reasons to accept it comport with our rejection of aggregation.

Our final commitment, common to many contractualists, is that one ought to minimize the maximum residual complaint of any individual. Together, our commitments imply:

> Rank-Matching Theory
>
> One ought to minimize the maximum residual complaint relative to any alternative once the person with the k-th strongest interests in favor of the alternative considers the person with the k-th strongest interests in favor of the option as justification for that option.

We imagine a hypothetical sequence of justifications where the person with the most to gain from some option considers the interests of the person with the most to gain from the alternative. This

[14] This feature distinguishes our view from the balancing view in Kamm 1984: 182, Kamm 1993: 114 – 119, and Kamm 2006: 57 – 61. See also Kumar 2001 and Kumar 2011: 140 – 143.

weakens, or even eliminates, any residual complaint of the former. Similarly for whoever has the second-most to gain from each option, the third-most, and so on. Each person is matched with a counterpart of equal rank on the other side, so that so far as possible, stronger interests are matched with stronger and weaker with weaker. After each person considers the perspective of their counterpart and modifies their assessment of the options accordingly, we heed the strongest residual complaint.

2.1 Justification and Residual Complaint

The core idea, that we can justify an action to someone by asking them to consider the perspective of another, should be familiar. To illustrate, suppose that you are attending a talk and two friends, Alice and Bob, both ask to borrow a pencil. Rummaging in your bag, you only find one. It is permissible to lend the pencil to either friend. If, say, Alice protests that your lending the pencil to Bob leaves her unable to write, you could reply that lending the pencil to Alice leaves Bob in just the same position. Each can be asked to consider the perspective of the other as justification for helping the other, and after doing so, neither can appropriately complain.

Conceptually, we distinguish three stages:

1. Alice and Bob each have an interest in getting the pencil.
2. Alice considers Bob's interests and vice-versa.
3. Neither Alice nor Bob can complain if you give the pencil to the other.

The complaints (if any) that remain at the final stage are called *residual complaints*, and they determine the permissibility of an action on our view. To be clear, residual complaints are normative objects that do not depend on whether someone actually complains, raises their tone, or leaves a negative review. Relatedly, the role of consideration is hypothetical: it is permissible to give the pencil to Bob whether or not Alice actually considers Bob's interests. What matters is the availability of such justifications and the residual complaints they leave people with.

We claim that these residual complaints come in degrees, and that the strength of a residual complaint corresponds to the difference in the strength of the interests of the relevant parties. To bolster this claim, consider some variations that strengthen Alice's interests and weaken Bob's.

Suppose that Alice needs the pencil to take notes, while Bob just wants to doodle. Again, we could ask Alice to consider Bob's desire to doodle. But now it seems appropriate for Alice to complain if we choose Bob over her. If we strengthen Alice's interests further (e.g. she needs the pencil for an exam), this residual complaint seems all the stronger. On the flip-side, by strengthening Bob's interests, we can weaken Alice's residual complaint. If Bob needs the pencil to fill out financial aid paperwork that is due within the hour, her residual complaint may be extinguished.

These patterns suggest that the strength of Alice's residual complaint corresponds to the degree by which her interests exceed those of Bob. When Bob's interests are stronger, she has no residual complaint. When her interests are stronger, she has a residual complaint whose strength depends on the degree of excess. To compare the residual complaints of two people, then, we need comparisons like those in Interpersonal Difference of Interests.

We note a limiting case: sometimes one has no justification to offer. Suppose you decline to lend your pencil to Alice for no reason at all. Here, Alice's interests are offset by nothing whatsoever, so the strength of her residual complaint corresponds to the strength of her interests. As we will see in §3, this case arises in our theory whenever the number of people who stand to gain from one option exceeds the number who stand to gain from the alternative.

2.2 No Repetition

A key principle in our theory is that the same person's interests cannot weaken or quell the complaints of multiple others. To see what this means, imagine that Alice, Bob, Carol, David, Edie, and Frank must decide where to go for dinner. The candidates have been winnowed to two: The Admiral, a local pub (A), or Boqueria, a tapas bar (B). Alice prefers A, but everyone else prefers B. Suppose that you are tasked with deciding where to go. Everyone trusts your just and dispassionate temperament. Would it be reasonable to choose A?

Here is a way to make the case for A. We could ask Bob, who prefers B, to consider Alice's similar preference for A as justification for choosing the latter. This would leave Bob with no residual complaint. Then we could ask Carol, who prefers B, to consider Alice's similar preference. This would leave Carol with no residual complaint. And so on through all the rest, leaving none with any residual complaint against choosing A. Since no one has any remaining complaint, there is nothing wrong with choosing A.

This reasoning, we submit, sounds terrible. It is double-counting to allow Alice's interests to answer, not just Bob's competing interests, but also Carol's competing interests and so on, thereby quelling any residual complaint. It treats this choice between one and five as though it were between five and five. Intuitively, Alice's interests cannot be repeatedly deployed to weaken or quell the residual complaints of multiple others. More generally, we accept:

> No Repetition
>
> The same person's interests cannot weaken or quell the residual complaints of multiple others.

We see no good argument for the contrary principle that allows the same person's interests to be repeatedly deployed as above. But we anticipate a concern arising: how can No Repetition comport with our rejection of aggregation? Doesn't sensitivity to the numbers imply an aggregative worldview? The strongest version of this concern asserts that No Repetition is logically inconsistent with our non-aggregative commitments. This claim is unfounded: No Repetition does not require resources that go beyond the non-aggregative domain. But there are two weaker forms of the concern worth addressing:

- Motivational Conflict: The only good reasons to accept No Repetition conflict with our reasons for accepting No Sums of Interests.

- Theoretical Collapse: No Repetition leads to a theory that differs little from aggregative moral theories, voiding the attempt to develop a non-aggregative moral theory of interest.

Both concerns gain plausibility from a more fundamental error worth first addressing: conflating the question of whether numbers have any moral significance with the question of whether the sum of people's interests has any moral significance. The two questions are related since an aggregative moral theory *must* give the numbers moral significance: by increasing the number affected, one increases the sum of their interests. But it does not follow (and is not true) that non-aggregative moral theories must give the numbers *no* moral significance.

To see this, consider the view that one should maximize the sum of the number of people that one helps and the number of people that lead happy lives. Suppose that this is claimed to be a fundamental moral truth, not susceptible of further justification or explanation. This theory is compatible with our non-aggregative commitments: it uses only information about whether one helps some individual and whether some individual leads a happy life. It sums the *number* of people helped with the *number* of people who will lead happy lives, but non-aggregative moral theories do not claim (implausibly) that *numbers* do not sum.

Despite their entanglement, the question of whether the numbers count and whether the interests of separate people sum are importantly distinct. We reject aggregation on the basis of a metaphysical picture that analogizes the interests of separate people to the beauty of separate paintings. This is a view about what facts are available for use in moral theory or elsewhere; theories cannot use "facts" that do not obtain.[15] Meanwhile, our reason for accepting No Repetition appeals to the thought that it is unfair to allow one to silence so many. We do not rely on the thought, which we regard as mistaken, that the many collectively have more at stake. Rather, the intuition is just that this procedure seem unbalanced, disproportionate, unreasonable. Each counts for one. That is why one cannot count for many.

None of this endangers the thought that the borders between the lives of separate people are permanent fixtures in the moral landscape, and that one cannot pool the interests of separate

[15] Distinguish this from the acceptable practice of using the fact that some proposition does not obtain.

people across these lines. This addresses the Motivational Conflict. We defer the worry about Theoretical Collapse to §3, where it will be clear that our theory has implications that differ in striking ways from aggregative, as well as other non-aggregative, moral theories.

2.3 Minimax Residual Complaint and Rank-Matching Theory

In principle, many ways of asking individuals to consider one another's interests comport with No Repetition. In our restaurant example, Alice's interests could quell Bob's residual complaint or Carol's. No Repetition requires only that her interests do not quell both. Let a *matching* be any specification of who considers whose interests that satisfies No Repetition. The possible matchings are constrained, since no one may be matched with more than one person, but not unique. Moreover, even after fixing the matching, there is a question about how residual complaints determine what we ought to do. We resolve both ambiguities with the following principle:

> Minimax Residual Complaint
>
> One may choose an option iff for any alternative, there exists a matching such that the maximum residual complaint against choosing this option over the alternative does not exceed the minimum possible maximum residual complaint relative to any alternative.

Let us unpack this principle. Given a matching, we can determine the strength of an individual's residual complaint against choosing some option over another. We can then compare these individuals' residual complaints to determine whose is strongest.[16] This is the maximum residual complaint of any individual against choosing that option option over the other relative to our matching.

The maximum residual complaint depends on the matching. For example, the empty matching (which fails to match anyone with anyone) generates higher maximum residual complaints than many alternative matchings. Minimax Residual Complaint considers the minimum possible maximum residual complaint across all options and alternatives and all possible matchings.

[16] Ties can be broken arbitrarily.

Think of this as the theoretical minimum for the maximum residual complaint. It then permits those options that, relative to some matching for any alternative, yield a maximum residual complaint no greater than the theoretical minimum. The idea is that these options do as good a job as possible, given the situation, to minimize the strongest individual complaints.

Given the strengths of people's interests and residual complaints, Minimax Residual Complaint settles which options are permissible in any choice. Furthermore, it implies our theory:

Fact 1. One may choose an option according to Minimax Residual Complaint iff one may choose that option according to the Rank-Matching Theory.

We defer the details to the appendix, but the main idea is simple. Given two options, let *rank-matching* name any matching that pairs someone with the k-th strongest interests in favor of one option to someone with the k-th strongest interests in favor of the alternative.[17] It turns out that rank-matching minimizes the maximum residual complaint for any option relative to any alternative. So to determine which option achieves the minimum possible maximum residual complaint (as Minimax Residual Complaint demands), we can simply consider the maximum residual complaint against each option in any corresponding rank-matching (as Rank-Matching Theory demands). Thus, as promised:

> Rank-Matching Theory
>
> One ought to minimize the maximum residual complaint relative to any alternative under any rank-matching.

Before turning to the next section, it is worth highlighting the forest in these trees. We ask individuals with the strongest interests to consider the strongest competing interests, and we ask those with weaker interests to consider weaker competing interests. We do so in an effort to make our choice as acceptable as possible to the person to whom it is least acceptable. These are not technical or obscure ideas, even if they can lead to technical places.

[17] There can be multiple rank-matchings for a given pair of options. As the appendix shows, these rank-matchings agree on the maximum residual complaint relative to any alternative, so the prescriptions of the Rank-Matching Theory remain well-defined (Corollary 7).

3 Synchronic Implications

Our primary case for the Rank-Matching Theory is that it follows from the preceding (non-aggregative) picture of the interests of separate people when coupled with No Repetition and Minimax Residual Complaint. Fact 1 and Fact 4 in the appendix make this precise.

Our secondary case appeals to the theory's implications. Agents face many decisions over their lifetimes. In light of this, there are two ways to think about the implications of the Rank-Matching Theory for a particular choice.[18] On the synchronic approach, what one ought to do depends solely on one's present options, considered in isolation from any other choices one may face.[19] On the diachronic approach, what one ought to do depends also on the other choices one has made or will make.[20]

We begin with the implications of the Synchronic Rank-Matching Theory, which applies the Rank-Matching Theory to individual choices considered in isolation from any others. This has both plausible and implausible implications that other views lack. This may seem a damaging result, but in the next section, we contend that the Rank-Matching Theory should anyways be applied diachronically, taking into account one's other decisions. And we show how doing so, at least in one simple model, leads the Rank-Matching Theory away from implausibility.

3.1 Shared Implications of Synchronic Rank-Matching Theory

We begin with some implications of the Synchronic Rank-Matching Theory that rival views share. We walk through examples to illustrate why these implications obtain and how the theory works. The proofs, which are all straightforward, are in the appendix.[21] We begin with:

Disparate Interests

[18] We set aside other ways, such as naive and sophisticated choice, for reasons of space.

[19] This is also known as myopic choice, which differs from naive choice. In naive choice, one considers and plans for future choices without considering how one may deviate from those plans (Gustafsson 2022: 6).

[20] This is also known as resolute choice (McClennen 1990: 13).

[21] See §A.1.

If someone has interests that are far stronger than anyone's competing interests, then all else equal, one ought to help them.

To see how the Rank-Matching Theory satisfies this principle, recall Transmitter Room. Rank-matching pairs Jones with whichever audience member has the most at stake. All other audience members go unmatched. Even after considering the interests of this audience member, Jones retains a very strong residual complaint against allowing him to suffer. None of the other audience members have anything like a comparably strong interest in avoiding the delay, so helping Jones minimizes the maximum residual complaint under this rank-matching. Thus, we ought to do so.

Another shared implication is:

<u>Equal Interests</u>

If each affected party has equally strong interests at stake, then all else equal, one ought to help the greater number.

Recall our restaurant example, where we can either go to The Admiral (as Alice prefers) or Boqueria (as Bob, Carol, David, Edie, and Frank prefer). Our rank-matching pairs Alice with, say, Bob. This leaves neither with any residual complaint. But Carol, David, Edie, and Frank may still complain about going to The Abbey. Thus, going to Boqueria minimizes the maximum residual complaint under this rank-matching. So we ought to do so.

A third shared implication is:

<u>Tradeoffs</u>

For any strength of interest S, there exists a strength of interest T and some number n such that all else equal:

(i) one ought to help one person with interests of strength S rather than another with interests of strength T, and

(ii) one ought to help n people with interests of strength T rather than one person with interests of strength S.

This implication is noteworthy because it distinguishes the Synchronic Rank-Matching Theory from a version of leximin that requires one to, say, lexically minimize the strongest unsatisfied interest. Disparate Interests shows that the strength of an interest can matter on this theory, while Equal Interests shows that the numbers can also matter. Tradeoffs adds that changes in the numbers can sometimes override differences in the strengths of competing interests.

To illustrate how the theory satisfies this principle, we consider the following prescriptions (which exemplify Tradeoffs):

- One should provide Alice with health and dental insurance (S) rather than provide Carol with health insurance (T).
- One should provide health insurance (T) to Bob and Carol rather than provide Alice with health and dental insurance (S).

In the first choice, rank-matching pairs Alice and Carol. This extinguishes Carol's residual complaint against helping Alice but leaves Alice with some residual complaint against helping Carol. Thus, we ought to help Alice.

In the second choice, rank-matching pairs Alice with, say, Bob. This leaves Bob with no complaint against helping Alice, but Carol retains a fairly strong residual complaint against doing so. Having considered Bob's interests, Alice's residual complaint is weaker than Carol's residual complaint. Thus, we ought to help Bob and Carol.

3.2 Distinctive Implications of Synchronic Rank-Matching Theory

We turn to some implications of the Synchronic Rank-Matching Theory that rival views lack. Consider:

> Nearly Equal Interests
>
> For any interest, there exist stronger and weaker variants such that: if everyone's interests lie within their range, then all else equal, one ought to help the greater number.

To see the appeal of Nearly Equal Interests, consider:

> Islands
>
> There are a zillion people stranded on Island A and a zillion and one people stranded on Island B. You can rescue either but not both groups. Whichever group you do not rescue will perish.
>
> As it happens, the winds favor the former group: it would take six days to reach Island A, but it would take seven days to reach Island B. This difference will not affect the survival of anyone on either island; they all have enough food and supplies to last a week. But the time spent stranded is hardly pleasant; people are bored and anxious and stressed. It is better to be rescued sooner. What should you do?

Intuitively, one ought to head to Island B. Although each person on Island A may benefit slightly more (from the slightly earlier rescue) than those on Island B, going to Island B will save an additional life. This latter fact seems decisive, especially on a non-aggregative theory that declines to sum benefits across people, and it is captured by Nearly Equal Interests.

Let us see how the Synchronic Rank-Matching Theory satisfies Nearly Equal Interests in this case. The rank-matching pairs a zillion of those on Island A with a zillion on Island B. This nearly extinguishes the residual complaint of each individual on Island A. But someone on Island B goes unpaired and they retain a very strong residual complaint. Saving those on Island B thus minimizes the maximum residual complaint and so we ought to do so.[22]

The Synchronic Rank-Matching Theory also has a distinctively counter-intuitive implication:

> Nearly Innumerate
>
> One may (must) help someone with interests of strength S rather than two people with interests of strength T iff one may (must) help someone with interests of strength S rather than n people with interests of strength T, provided n is at least two.

[22] Astute readers will note that the requirement to help those on Island B goes through even when the zillion-and-first person on Island B has far less at stake, which may be counter-intuitive. We take this to be a milder version of the problem that Nearly Innumerate raises, so we do not address it separately.

This implies that we must help one person with interests S rather than two with interests T iff we must help one person with S rather than, say, a million with T. For many choices of S and T, this is extremely implausible. Here is one example: we must provide one person with a life-saving appendectomy rather than prevent two people from going blind iff we must provide one person with the appendectomy rather than prevent a million people from going blind.

This principle makes the Synchronic Rank-Matching Theory nearly insensitive to the numbers. We say "nearly" because the difference between one and two remains.[23] The principle ties together all comparisons between S and two or more instances of T, but it does not constrain how S compares to a single instance of T. Moreover, that the theory distinguishes one from two can be seen in the preceding discussion of Tradeoffs. The Synchronic Rank-Matching Theory is therefore sensitive to the difference between one and two, but not between two and anything greater.

To see why it has this feature, consider the choice between saving Alice's life and sparing Bob and Carol from blindness. We assume that Bob and Carol have equally strong interests at stake, and our rank-matching pairs Alice with, say, Bob. This diminishes Alice's residual complaint and extinguishes Bob's but it leaves Carol's residual complaint unaffected. The key question is whether Alice's residual complaint exceeds Carol's in strength or not. If it does, then we must save Alice. If it is equally strong, then we may help either group. If it is weaker, then we must help Bob and Carol.

Now imagine the choice between saving Alice's life and sparing not just Bob and Carol but also David, Edie, ..., Zebediah from blindness. Like before, we assume that Bob, Carol, ..., Zebediah have equally strong interests at stake and so our rank-matching pairs Alice with, say, Bob. This diminishes Alice's residual complaint and extinguishes Bob's but it leaves Carol's, David's, ..., Zebediah's residual complaints unaffected. The key question is whether Alice's residual complaint exceeds that of Carol, David, ..., Zebediah. Since the interests of the latter are equally

[23] Along with, of course, the difference between zero and one.

strong, this is equivalent to asking whether Alice's residual complaint exceeds, equals, or falls short of Carol's residual complaint. Alice's residual complaint will be stronger than Carol's iff it is stronger than David's iff … iff it is stronger than Zebediah's. So we may (or must) help Alice in this choice iff we may (or must) help Alice in the preceding choice.

This feature arises because we do not aggregate the interests or residual complaints of separate people. We are therefore left with a series of ordinal comparisons between Alice's residual complaint, Carol's residual complaint, David's residual complaint, and so on. The only sensible rule then seems to be something like maximin, requiring us to help Alice when her residual complaint is strongest.

For what it is worth, we suspect that any non-aggregative moral theory will face some version of this problem. The thought is that any such view must at bottom have some morally significant property whose tokens can only be ordinally compared.[24] As a result, the final rule must be something like maximin or headcount (which requires that one minimize the number of tokens above some threshold).[25] If indeed these are the only options, then this problem arises. The issue is not exactly that non-aggregative views must be insensitive to the numbers; as §2.2 argues and headcount exemplifies, this is not true. Rather, the problem is that non-aggregative views cannot smoothly integrate differences in degree with differences in quantity into an overall ranking. Such tradeoffs among dimensions, we suspect, will have some jagged edge on any non-aggregative view. But making this hunch precise must be left for future work.

[24]Why not take differences of these tokens and compare these differences, as our view does when considering the interests of separate people? The worry is that allowing such differences to be taken indefinitely leads to an aggregative structure, as §1.2 and Footnote 12 discuss. It is reasonable, then, to surmise that the process of taking differences must end at some level whereupon this problem will arise.

[25]See Nebel 2024: 297–298 for reasons to think headcount compatible with merely ordinal comparisons.

4 Diachronic Implications

These implications result from applying the Rank-Matching Theory to single decisions considered in isolation. But that is a strange way to apply the theory. For one, it is not what we ordinarily do. We generally think in terms of decisions (e.g. to eat a sandwich, to get a job, to raise a child) that, strictly speaking, involve a string of individual decisions. And the kinds of entities in our world that largely determine the availability of services like appendectomies or cataract surgeries, such as government agencies and hospitals and insurance companies, also seem to make decisions that take into account their past and future choices. Doing so seems sensible, while the idea that we would only consider the effects of the present choice seems bizarre. Moreover, the effects of the present choice considered in isolation are rarely well-defined. When, as is common, one's present choices shape one's future options, it is not clear what counts as the effect of one's present choice, considered in isolation from what one will choose in the future.

For another, what other choices we face seems to affect what we should do. The imperative to rescue Jones in Transmitter Room seems so clear partly because such choices arise so rarely.[26] If they arose regularly, we might well judge the case differently. Imagine, for example, that such incidents were routine and that a general policy of shutting down transmissions whenever necessary to save someone from an hour of extremely painful electrical shocks would mean, in practice, that transmissions are constantly interrupted, nobody bothers trying to watch anything, no firm attempts to transmit content, and these technologies go unused. If this were the only alternative given the technological constraints, we might well tolerate these accidents as the necessary cost of broadcasting just as people tolerated deaths from the introduction of motor vehicles. For those with doubts, it may help to consider a parallel case. If one can save a drowning child at the expense of a few thousand dollars (via, e.g., a ruined suit), few doubt that one must do so. But the rarity of the case seems to matter. If such cases arose regularly, as may be true for many in

[26] Scanlon 1998: 238.

high-income countries, one's judgment about the obligation to help in any particular case would probably change well before becoming destitute.[27] Here too, the presence of these other choices makes an intuitive difference.

These concerns provide reason to apply the Rank-Matching Theory diachronically, considering a series of decisions rather than a single choice in isolation. How, if at all, does this change the resulting implications? To answer this question, we need a model of the choices in the series.

4.1 Model

Imagine a population whose members begin at the same initial level of wellbeing. An agent faces a series of binary choices between maintaining the status quo and opting for some change. The final outcome of any series of choices is determined by summing the changes that the agent accepts. To avoid the complications of uncertainty, we assume that the agent knows exactly what results from any choice or sequence of choices.[28]

Our model includes two kinds of choices: background and focal. In background choices, the effects of opting for change at any particular choice are determined at the outset by a given probability distribution (metaphorically: chosen by Nature). Background choices are meant as a reasonable, if strong, idealization of the kinds of choices that policymakers often face. The precise effects are to some degree random because how particular choices affect particular individuals depends on features of the world that vary from choice to choice. But the effects are generated by a common distribution because, we assume, these choices tend to share a common shape. For example, exactly who gets injured and thus receives workers compensation may vary across nearby worlds, but the number of workers that benefit may be fairly robust.

[27] See Timmerman 2015 and chapter six of Pummer 2023.

[28] The assumption of such knowledge is unrealistic, but it is an unrealistic assumption that we have been making throughout. We do so because there is substantial disagreement about how to extend contractualist theories, of which ours is an example, to risky cases (see, e.g. Fleurbaey and Voorhoeve 2013, Frick 2013, Lazar 2018, Otsuka 2015). Resolving this question would take us too far afield. Moreover, the implication that worries us, Nearly Innumerate, arises even in the riskless setting. It is therefore worth considering whether it can be ameliorated in this setting.

In a focal choice, by contrast, the available change is fixed by us in order to test the implications of the Rank-Matching Theory. We consider how the presence of the background choice(s) changes what the theory prescribes for the focal choice. Our model thus has several layers of background choices before a final focal choice:[29]

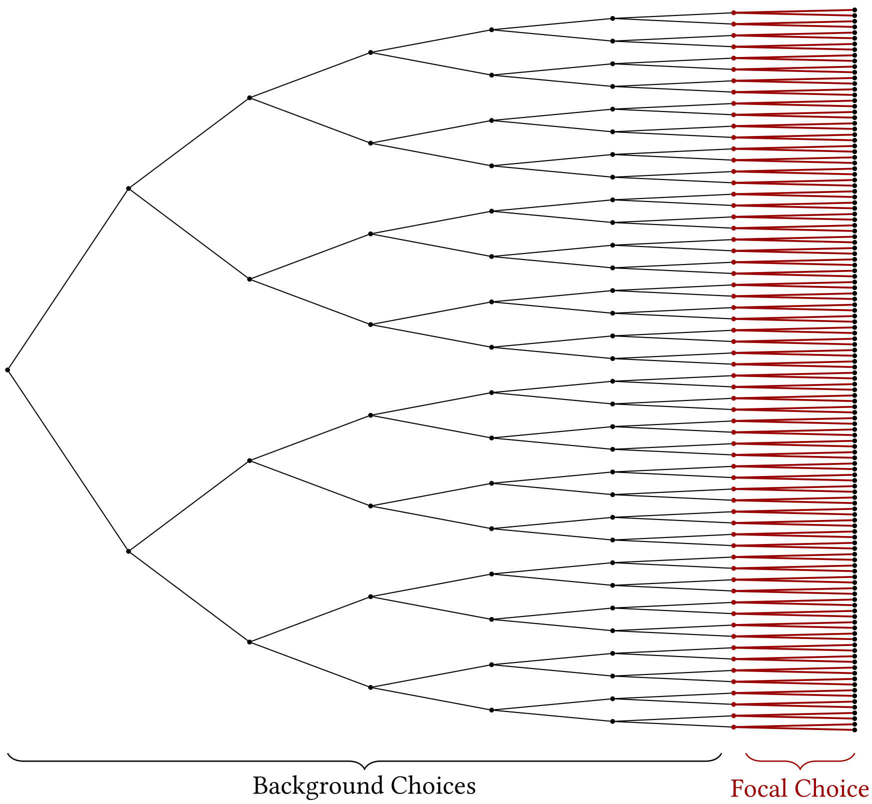

Figure 1: The model

What are the changes available at background choices? We assume that these changes are independent and identically distributed. 'Independent' means that the change available at one background choice is statistically independent of the change(s) available at any other. 'Identically distributed' means that the probability that any change is available at any background choice is the same. These assumptions are unrealistic: prior decisions influence later opportunities, and

[29]We fix the focal choice to be the final choice because when the focal choice comes earlier in the sequence, it has two separate effects: (i) the effect of the choice itself, and (ii) the effect of the choice on which downstream alternatives become available. We want to isolate the former here.

the profile of opportunities varies across different background choices. But relaxing them would substantially complicate the model, so we adopt them as simplifying assumptions. If the model proves fruitful, future work can explore weaker assumptions.

We further assume that the change option at each background choice determines the wellbeing change for each person by an independent draw from a normal distribution. This means that in any background choice, changes to one person's wellbeing are uncorrelated with changes to anyone else's wellbeing. Again, this is not realistic; if Alice and Bob earn high incomes while Carol earns a low income, then the effects of changes on Alice and Bob may be strongly correlated while the effects on Alice and Carol may be weakly or even negatively correlated. It is natural to modify this assumption by introducing groups with different intra-group and inter-group correlations. Again, these modifications can be explored in future work.

We stress that these probability distributions are used only to generate the decision tree, which happens before the agent makes any decisions. They represent variation in the state of the world, rather than uncertainty on the part of the agent. The agent always knows exactly what results from any choice or sequence of choices through the tree.

4.2 Applying the Rank-Matching Theory

How do we apply the Rank-Matching Theory? So far we have only considered cases where the agent makes one choice from two options. What happens when the agent faces a series of choices?

When the agent knows what results from any (complete) series of choices, we imagine them choosing directly from the set of possible final outcomes. Some option(s) will be permissible. Then, we permit the agent to make any series of choices that leads to one of these outcomes. To illustrate, consider the following choice:

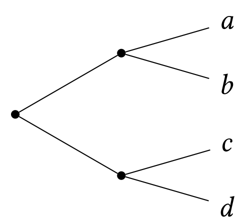

Figure 2: Two binary choices

To determine what one may choose, we consider what one could permissibly choose when directly selecting among the final outcomes. If, say, one ought to choose c, then one ought to choose a path that leads to c:

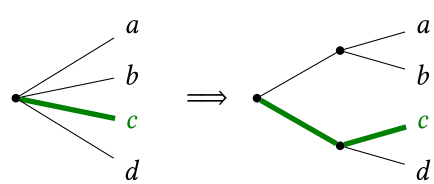

Figure 3: The reduction

This reduces the problem of making a series of choices to the problem of making a single choice from many options. But how should the agent choose when presented with many options?

Like before, Minimax Residual Complaint settles this question. Recall:

> Minimax Residual Complaint
>
> One may choose an option iff for any alternative, there exists a matching such that the maximum residual complaint against choosing this option over the alternative does not exceed the minimum possible maximum residual complaint relative to any alternative.

With more than two options, we can think of this principle as follows. For each option, consider the possible alternatives. For each comparison, choose the matching that minimizes the maximum residual complaint against choosing the preceding option over the alternative.[30] Then, consider the strongest residual complaint that arises in any such comparison between the option and some

[30] Again, ties can be broken arbitrarily.

alternative. We could call this the unavoidable complaint that the option raises. Minimax Residual Complaint permits choosing an option iff that option minimizes the unavoidable complaint.

As before, rank-matching plays a special role. Recall:

> Rank-Matching Theory
>
> One ought to minimize the maximum residual complaint relative to any alternative under any rank-matching.

With more than two options, we can think of this principle as follows. For each option, consider the possible alternatives. For each comparison between an option and some alternative, choose a rank-matching. Then, consider the strongest residual complaint that arises against this option in any rank-matching with any alternative. We could call this the unavoidable rank-matched complaint that the option raises. The Rank-Matching Theory permits choosing an option iff that option minimizes the unavoidable rank-matched complaint. And as mentioned, the appendix proves:

Fact 1 (restated)**.** One may choose an option according to Minimax Residual Complaint iff one may choose that option according to the Rank-Matching Theory.

4.3 Results

We consider four simulations. Each has zero to six background choices and then a focal choice that varies:[31]

- Simulation 1 (moderate vs. mild): the focal choice is between raising Person 1's wellbeing by ten units while lowering some others' wellbeing by three units or doing nothing
- Simulation 2 (severe vs. moderate): the focal choice is between raising Person 1's wellbeing by fifty units while lowering some others' wellbeing by fifteen units or doing nothing

[31] The (anonymized) code for each simulation is here: https://figshare.com/s/eb95e5dd2ee8ed04b2f8.

- Simulation 3 (severe vs. mild): the focal choice is between raising Person 1's wellbeing by fifty units while lowering some others' wellbeing by one unit or doing nothing

- Simulation 4 (severe vs. mild): same as Simulation 3, except the population is much larger

Each simulation fixes the magnitude of the gain to the one and the loss to the others while varying the number of background choices and the number of others who suffer the loss. We can therefore use these simulations to test the hypothesis that adding background choices will lead the Rank-Matching Theory away from the troubling implications of Nearly Innumerate.

In any focal choice, Nearly Innumerate implies that one should accept the gain for Person 1 when doing so burdens two others iff one should accept the gain for Person 1 when doing so burdens, say, ninety-nine others. Changing the number of people who suffer the loss does not, according to Nearly Innumerate, change the prescriptions of the theory. And, indeed, this is what we find when we consider each focal choice in isolation (by setting the number of background choices to zero). Start with Simulation 1:

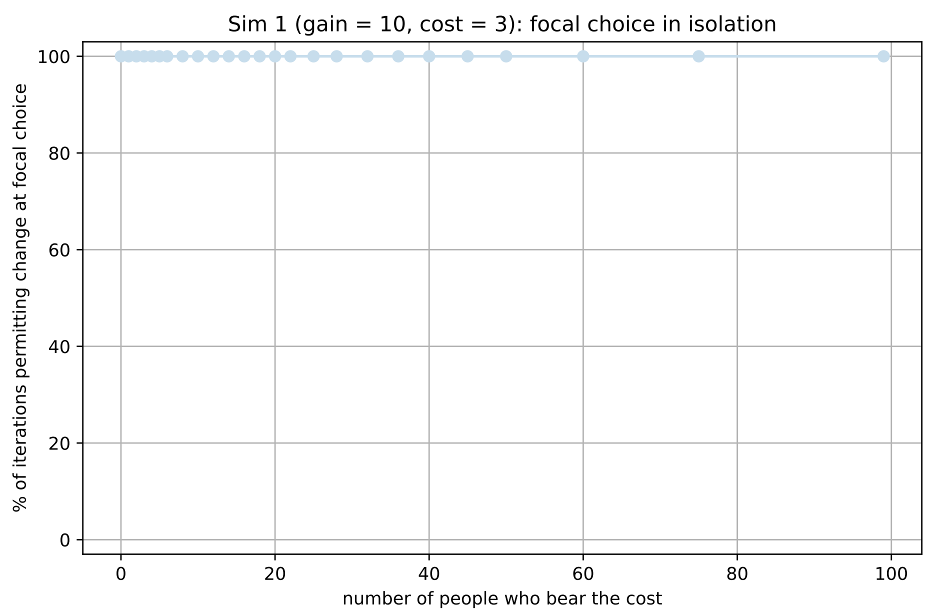

Figure 4: Simulation 1 without background choices.

Let me explain the chart. Because our simulations contain elements of randomness (from the background choices), we run a thousand iterations of each simulation. In Simulation 1, the focal choice is between increasing Person 1's wellbeing by ten units while lowering some number of others' wellbeing by three units or doing nothing. The horizontal axis varies the number of people whose wellbeing may be lowered by the focal choice. The vertical axis records the percentage of the thousand iterations where Rank-Matching Theory permits opting for change in the focal choice given the corresponding number of people who will have their wellbeing lowered. So the **(99, 100)** point in the top-right corner indicates that when there are no background choices, Rank-Matching Theory permits increasing Person 1's wellbeing by ten units while lowering **99** others' wellbeing by three units in **100%** of the thousand iterations we ran.

Figure 4 thus illustrates how, when we omit the background choices, the Rank-Matching Theory requires us to increase Person 1's wellbeing by 10 units irrespective of the number of others whose wellbeing will be lowered by 3 units. This accords with Nearly Innumerate. But what happens if we stop considering this choice in isolation and add some background choices?

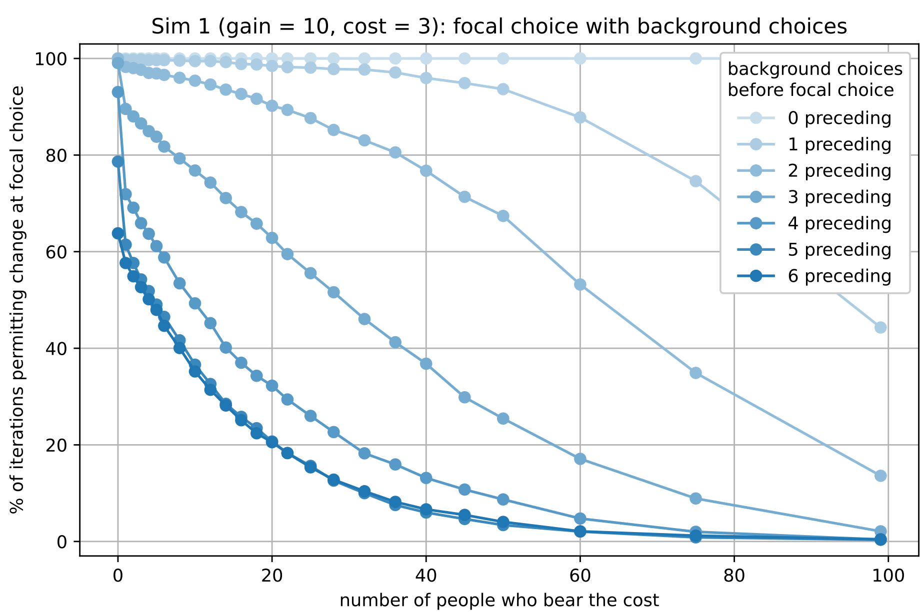

Figure 5: Simulation 1 with background choices.

The possibility of no background choices from Figure 4 is represented by the translucent line at the top. Iterations with more background choices are represented by more opaque lines. As we can see, *adding even a single background choice makes the number of people who bear the cost relevant to what one ought to do.* Moreover, this sensitivity grows as we increase the number of background choices. We thus see how the addition of such choices can lead the Rank-Matching Theory away from Nearly Innumerate and towards more plausible prescriptions.

Simulation 1 concerns a choice between the moderate interests of Person 1 and the mild interests of others. It exemplifies a tradeoff among relatively unimportant considerations. But similar remarks apply even after scaling the stakes as in Simulation 2:

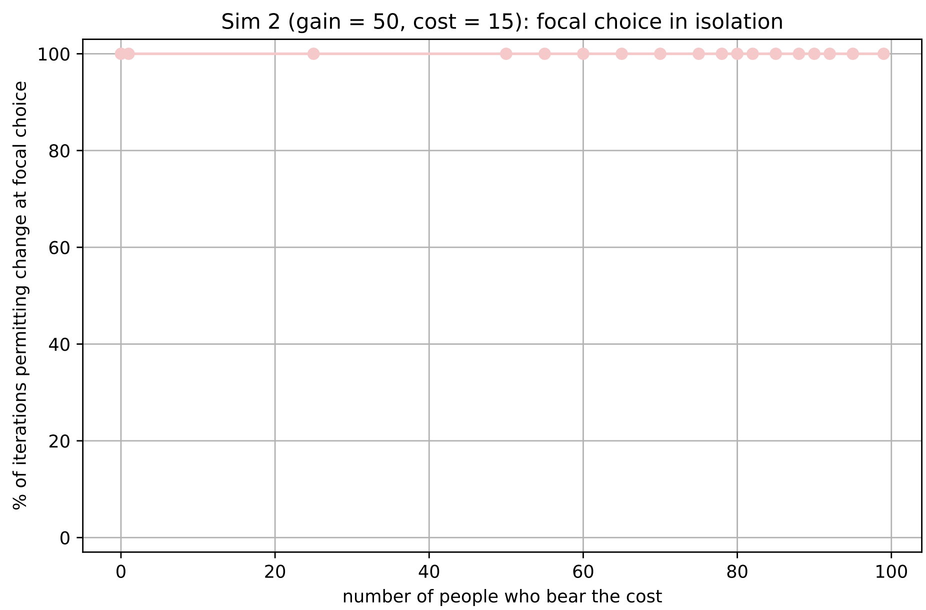

Figure 6: Simulation 2 without background choices.

Again, when we consider the focal choice in isolation, the number of cost-bearers does not affect what we may do. But when we add some background choices:

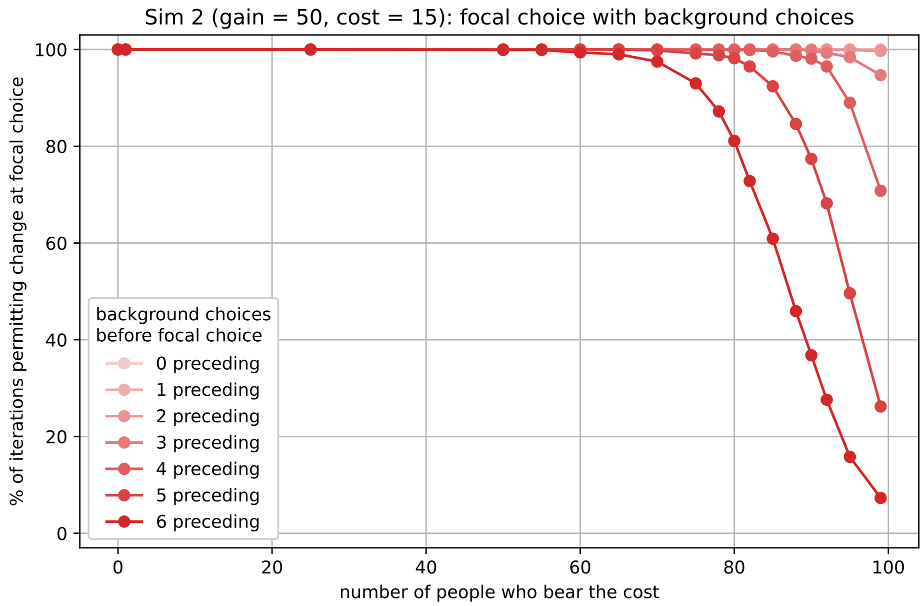

Figure 7: Simulation 2 with background choices.

We see that a similar pattern emerges.

Finally, we consider what happens when one person has far more at stake than anyone else. This tests whether the Rank-Matching Theory continues to satisfy Disparate Interests once we add these background choices. In Simulation 3, the focal choice is between increasing Person 1's wellbeing by fifty units while lowering some others' wellbeing by one unit and doing nothing. It resembles a choice between someone's severe interests and many others' mild interests. As before, when we consider this choice in isolation, we may help Person 1:

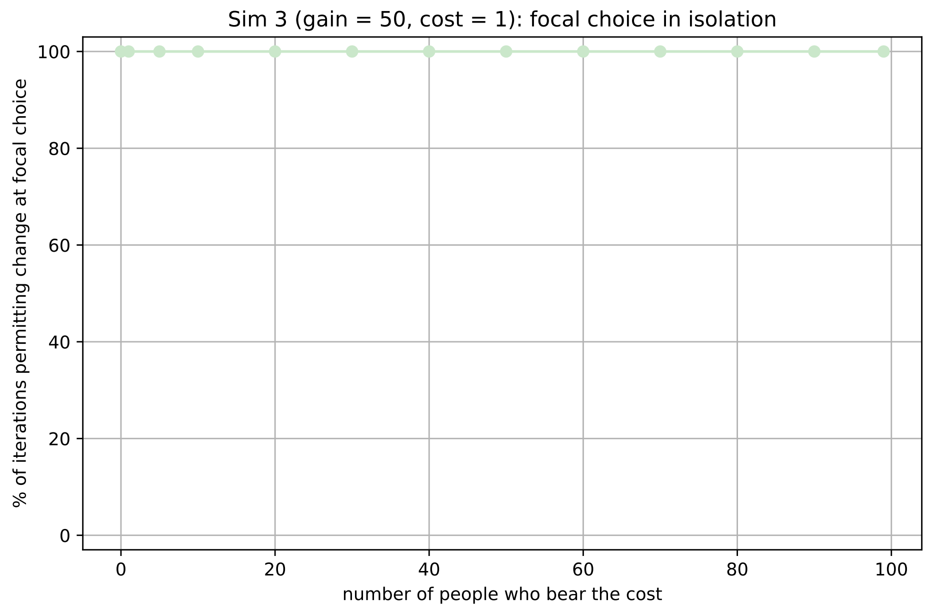

Figure 8: Simulation 3 without background choices.

But this time, when we add these background choices:

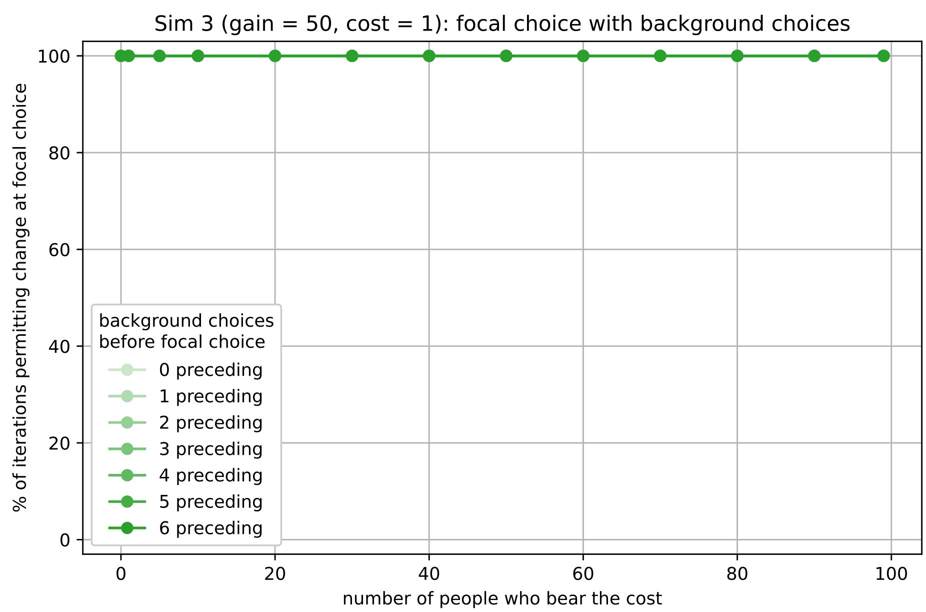

Figure 9: Simulation 3 with background choices.

We remain permitted to help Person 1 no matter the number of background choices.[32][33] Moreover, this remains true even after greatly increasing the number who bear the mild cost:

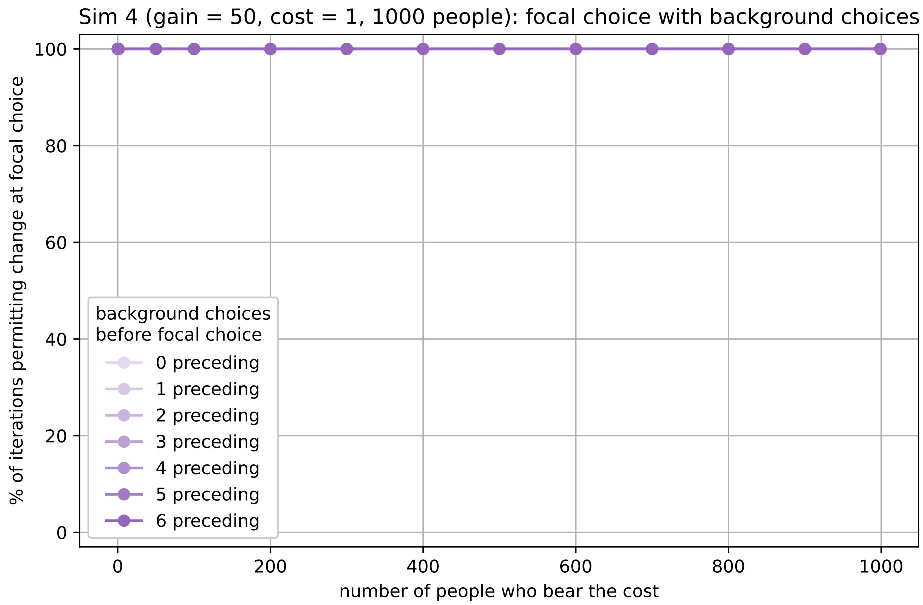

Figure 10: Simulation 4 with background choices.

One is tempted to say that Person 1's far stronger interests render irrelevant the competing interests of the many, so that one must help the former no matter how many there are of the latter. In this way, the Rank-Matching Theory may be able to provide a deeper and partly empirical account of why certain interests are relevant or irrelevant to others. The idea is that relations of relevance and irrelevance emerge from applying the Rank-Matching Theory to our circumstances, where importantly this includes the whole suite of decisions we face.

5 Conclusion

The Rank-Matching Theory follows from three related ideas. First, the interests of separate people do not sum. Second, one ought not double-count an individual's interests. Third, one ought to

[32]Moreover, inspection reveals that Rank-Matching Theory generically requires helping Person 1 in Simulation 3.
[33]We see one line rather than seven because they overlap.

minimize the maximum individual complaint. The first and second of these can be seen as a way to combine Taurek's critique of aggregation with sensitivity to the numbers. This might have seemed contradictory or theoretically underconstrained or little different from extant 'fully' or 'partially' aggregative views. Our investigation shows these not to be the case.

Moreover, the Rank-Matching Theory vindicates a number of plausible prescriptions. Its implausible prescriptions meanwhile look to be allayed by considering, as seems realistic, a backdrop of other choices. The result is a theory that prescribes sensitivity to the numbers generally but not when one person has far more at stake than anyone else. Future work may examine the robustness of these results and their applicability to the actual world.

We close with a final prediction of the theory. The foundations of the Rank-Matching Theory differ starkly from aggregative views like utilitarianism. But when we expand the suite of relevant decisions, its conclusions begin to resemble those of an aggregative theory. This, we have argued, makes the view more credible. But it also suggests that as we come to know more, control more, and coordinate more, it will become ever more appropriate to think of separate decisions as part of the same tree. Applying the Rank-Matching Theory to these richer trees may then yield conclusions that come ever closer to those of aggregative views. There thus seems to be a way in which merely expanding the sphere of action forces us to stomach tradeoffs which, on their own, we would not favor. If the Rank-Matching Theory is correct, then as our powers increase, we may have to accept ever more alienation from the particular choices we face and trust that the overall sequence is pointed in the right direction.[34]

[34] This contrasts with transitive theories, such as utilitarianism, that align local and global choice (Hammond 1988).

A Appendix

We state and prove the results of the paper.

$\mathbb{N} = \{1, 2, 3, ...\}$ represents the set of all *individuals*.

$\mathbb{R}$ represents the set of all lifetime *wellbeing levels*.

$\mathbb{R}^{\mathbb{N}} = \{f : \mathbb{N} \to \mathbb{R}\}$ represents the set of all *outcomes*.[35]

Definition 1 (Choices)**.** A *choice* is any finite subset $X \subseteq \mathbb{R}^{\mathbb{N}}$ such that:

$$|\{i \in \mathbb{N} : \exists x, y \in X : x_i \neq y_i\}| < \infty.$$

$\mathcal{X}$ is the set of all choices.

Choices so defined include only cases where agents have finitely many options and can affect the lifetime wellbeing of finitely many individuals.[36]

Definition 2 (Interests)**.** The strength of i's *interests* in x obtaining rather than y is:

$$s(x, y, i) = \max\{0,\ x_i - y_i\}.$$

We take the strength of i's interests to be the difference in i's wellbeing between the two outcomes, and it is zero when i fares no better under x than under y.

Definition 3. Let $n = |\{i \in \mathbb{N} : s(x, y, i) > 0\}|$ be the number of individuals with an interest in x over y. The *sorted interests* in x over y is $s(x, y) = (s_1, s_2, \ldots, s_n)$, where these n individuals' interests are arranged in weakly decreasing order: $s_1 \geq s_2 \geq \cdots \geq s_n > 0$.

Definition 4 (Matching)**.** The set of all *matchings* $\mathcal{M}$ is the set of all injections from $\mathbb{N}$ to $\mathbb{N}$. For any matching $\mu \in \mathcal{M}$, $\mu(i) = j$ means that individual i considers j's competing interests.

That any matching μ is injective implies that multiple individuals never consider the same person's competing interests, thereby satisfying No Repetition. That any matching μ is a function

[35] Outcomes are thus individuated by the lifetime wellbeing of each individual.

[36] We restrict the possible choices in order to avoid common problems with infinity.

implies that no individual ever considers the competing interests of multiple people, as No Sum of Interests requires.

Definition 5 (Residual Complaints)**.** Given a matching $\mu \in \mathcal{M}$, the strength of i's *residual complaint* against choosing y over x after considering their counterpart's interests is:

$$c_\mu(x, y, i) = \max\{0, s(x, y, i) - s(y, x, \mu(i))\}$$

which we also write as: $c_\mu(i : x \to y)$.

The strength of someone's residual complaint is the difference in the strength of their interests and the strength (which may be zero) of the interests they consider as justification.

Definition 2 and Definition 5 encode strong assumptions about the strengths of interests and residual complaints. We later show, in §A.2, that much weaker assumptions suffice for the equivalence between Minimax Residual Complaint and Rank-Matching Theory.

Let $x \in C(X)$ iff it is permissible to choose outcome x in choice X. Then:[37]

Principle 1 (Minimax Residual Complaint)**.** For any choice $X \in \mathcal{X}$:

$$C(X) = \arg\min_{x \in X} \max_{y \in X} \min_{\mu \in \mathcal{M}} \max_{i \in \mathbb{N}} c_\mu(i : y \to x).$$

Definition 6 (Rank-Matching)**.** Given options x and y, $\mu \in \mathcal{M}$ is a rank-matching iff

1. $\mu(\mathbb{N}) \supseteq \{i : s(y, x, i) > 0\}$, and

2. For any $i, j \in \mathbb{N}$: $s(x, y, i) > s(x, y, j) \implies s(y, x, \mu(i)) \geq s(y, x, \mu(j))$.

Let $RM(x, y)$ denote the set of all such rank-matchings.

As this definition suggests, in general there is no unique rank-matching. After all, multiple people may have interests of the same strength, and then rank-matching alone does not resolve who

[37] The more literal, but equivalent, formalization would be: $x \in C(X)$ iff $\forall y \in X, \exists \mu \in \mathcal{M}$:

$$\max_{i \in \mathbb{N}} c_\mu(i : y \to x) \leq \min_{x^* \in X} \max_{y^* \in X} \min_{\mu^* \in \mathcal{M}} \max_{i^* \in \mathbb{N}} c_{\mu^*}(i^* : y^* \to x^*)$$

gets matched with whom. But regardless of who exactly gets matched with whom, the strongest individual complaint against choosing one option over another is the same on any rank-matching:

Fact 2. For any $\{x, y\} \in \mathcal{X}$, $\mu \in RM(x, y)$, and $\rho \in \mathcal{M}$:

$$\max_{i \in \mathbb{N}} c_\mu(i : x \to y) \leq \max_{i \in \mathbb{N}} c_\rho(i : x \to y).$$

Proof. Let $c^*_\mu = \max_{i \in \mathbb{N}} c_\mu(i : x \to y)$ and let $c^*_\rho = \max_{i \in \mathbb{N}} c_\rho(i : x \to y)$. We show that $c^*_\mu \leq c^*_\rho$.

If $c^*_\mu = 0$, then we are done. So suppose that $c^*_\mu > 0$. Let i^* be such that $c_\mu(i^* : x \to y) = c^*_\mu$. Then, $c_\mu(i^* : x \to y) = s(x, y, i^*) - s(y, x, \mu(i^*))$. Let $a = s(x, y, i^*)$ and $b = s(y, x, \mu(i^*))$.

Since μ is a rank-matching, any k such that $s(y, x, k) > b$ has a corresponding j such that $\mu(j) = k$ and $s(x, y, j) \geq a$. Thus,

$$|\{i : s(x, y, i) \geq a\}| > |\{i : s(y, x, i) > b\}|.$$

Since ρ is injective, it cannot map the former into the latter. So there exists some $k^* \in \{i : s(x, y, i) \geq a\}$ such that $\rho(k^*) \notin \{i : s(y, x, i) > b\}$. So $c_\mu \leq c_\rho(k^* : x \to y) \leq c^*_\rho$. □

As a corollary, we obtain:

Corollary 7. *Let $\mu, \rho \in RM(x, y)$. Then,* $\max_{i \in \mathbb{N}} c_\mu(i : x \to y) = \max_{i \in \mathbb{N}} c_\rho(i : x \to y)$.

Proof. Switch the roles of μ and ρ in Fact 2. □

In light of this result, we can define:

Definition 8. The strength of the strongest residual complaint against choosing y over x on any rank-matching is:

$$c_{RM}(x, y) := \max_{i \in \mathbb{N}} c_\mu(i : x \to y)$$

where $\mu \in RM(x, y)$. Like before, we also write this as $c_{RM}(x \to y)$.

Principle 2 (Rank-Matching Theory)**.** For any choice $X \in \mathcal{X}$:

$$C(X) = \arg\min_{x \in X} \max_{y \in X} c_{RM}(y \to x).$$

In other words, the Rank-Matching Theory says to minimize the maximum residual complaint relative to any alternative under any rank-matching.

We can now show that Minimax Residual Complaint and Rank-Matching Theory are equivalent:

Fact 1 (restated)**.** For any choice $X \in \mathcal{X}$:

$$\arg\min_{x\in X}\max_{y\in X}\min_{\mu\in\mathcal{M}}\max_{i\in\mathbb{N}} c_\mu(i : y \to x) = \arg\min_{x\in X}\max_{y\in X} c_{RM}(y \to x).$$

Proof. It suffices to show that for any $\{x, y\} \in X \in \mathcal{X}$:

$$\min_{\mu\in\mathcal{M}}\max_{i\in\mathbb{N}} c_\mu(i : y \to x) = c_{RM}(y \to x)$$

Let $\rho \in RM(y, x)$. Since all rank-matchings are matchings:

$$\min_{\mu\in\mathcal{M}}\max_{i\in\mathbb{N}} c_\mu(i : y \to x) \leq \max_{i\in\mathbb{N}} c_\rho(i : y \to x) = c_{RM}(y \to x)$$

By Fact 2, for any $\mu \in \mathcal{M}$, we have:

$$c_{RM}(y \to x) = \max_{i\in\mathbb{N}} c_\rho(i : y \to x) \leq \max_{i\in\mathbb{N}} c_\mu(i : y \to x)$$

Thus, $c_{RM}(y \to x) \leq \min_{\mu\in\mathcal{M}} \max_{i\in\mathbb{N}} c_\mu(i : y \to x)$. □

A.1 Implications

We now turn to the implications of the Rank-Matching Theory. Let $x \succsim y$ iff $x \in C(x, y)$ and let $\succ$ and $\sim$ be defined in the usual way. Then:

Principle 3 (Disparate Interests)**.** There exist $s, t \in \mathbb{R}_{++}$ such that for any choice $\{x, y\} \in \mathcal{X}$:

$$\big(\exists i:\ s(x, y, i) \geq s\big) \ \wedge\ \big(\forall j:\ s(y, x, j) \leq t\big) \implies x \succ y.$$

Principle 4 (Equal Interests)**.** For any $s \in \mathbb{R}_{++}$ and any choice $\{x, y\} \in \mathcal{X}$:

$$\{s(x, y, i) : i \in \mathbb{N}\} \cup \{s(y, x, i) : i \in \mathbb{N}\} = \{0, s\}$$

implies

$$|\{i : s(x, y, i) > 0\}| > |\{i : s(y, x, i) > 0\}| \implies x \succ y.$$

Principle 5 (Tradeoffs)**.** For any $s \in \mathbb{R}_{++}$ there exists $t \in \mathbb{R}_{++}$ and $n \in \mathbb{N}$ such that:

1. $\exists \{x, y\} \in \mathcal{X} : s(x, y) = s \wedge s(y, x) = t \wedge x \succ y$, and

2. $\exists \{z, w\} \in \mathcal{X} : s(z, w) = s \wedge s(w, z) = \underbrace{(t, t, \ldots, t)}_{n} \wedge w \succ z$.

Principle 6 (Nearly Equal Interests)**.** For all $s \in \mathbb{R}_{++}$, there exist $s^+ > s > s^- > 0$ such that for all $\{x, y\} \in \mathcal{X}$:

$$\{s(x, y, i) \mid i \in \mathbb{N}\} \cup \{s(y, x, i) \mid i \in \mathbb{N}\} \subseteq \{0\} \cup [s^-, s^+]$$

implies

$$|\{i : s(x, y, i) > 0\}| > |\{i : s(y, x, i) > 0\}| \implies x \succ y$$

Principle 7 (Nearly Innumerate)**.** For any $s, t \in \mathbb{R}_{++}$ and $n \geq 2$, the following are equivalent:

1. $\forall \{x, y\} \in \mathcal{X} : s(x, y) = s \wedge s(y, x) = (t, t) \implies x \succsim y$

2. $\forall \{z, w\} \in \mathcal{X} : s(z, w) = s \wedge s(w, z) = \underbrace{(t, t, ..., t)}_{n} \implies z \succsim w$

and likewise with $\succsim$ replaced throughout by $\succ$.

Fact 3. The Rank-Matching Theory implies Disparate Interests, Equal Interests, Tradeoffs, Nearly Equal Interests, and Nearly Innumerate.

Proof. To show that $\{x\} = \arg\min_{a \in \{x,y\}} \max_{b \in \{x,y\}} c_{RM}(b \to a)$, it suffices to show that $c_{RM}(x \to y) > c_{RM}(y \to x)$. Given this observation, we consider each implication in turn.

Disparate Interests: Let $s = 3$ and $t = 1$. Let $\{x, y\} \in \mathcal{X}$ be such that:

$$\exists i : s(x, y, i) \geq s \wedge \forall j : s(y, x, j) \leq t$$

Let μ be any matching. Then, $s(y, x, \mu(i)) \leq 1$ and so $c_\mu(i : x \to y) \geq 2$ while $c_\mu(j : y \to x) \leq 1$. Thus, $c_{RM}(x \to y) \geq 2 > 1 \geq c_{RM}(y \to x)$.

Equal Interests: Let $s \in \mathbb{R}_{++}$ and let $\{x, y\} \in \mathcal{X}$ be such that $\{s(x, y, i) : i \in \mathbb{N}\} \cup \{s(y, x, i) : i \in \mathbb{N}\} = \{0, s\}$. Suppose $|\{i : s(x, y, i) > 0\}| > |\{i : s(y, x, i) > 0\}|$. Then, for any matching μ, there exists some i such that: $s(x, y, i) = s \wedge s(y, x, \mu(i)) = 0$ and thus $c_{RM}(x \to y) = s$. On any rank-matching $\rho \in RM(y,x)$ and for any individual j, $s(y,x,j) = s$ implies $s(x,y,\rho(j)) = s$ and thus $c_{RM}(y \to x) = 0$ as desired.

Tradeoffs: Let $s \in \mathbb{R}_{++}$, $t = \frac{3s}{4}$, and $n = 2$. Then, let $x = (s,0,0,0,...)$ and $y = (0,t,0,0,0,...)$. Then, $\{x,y\} \in \mathcal{X}$ and under any rank-matching s is paired with t. So $c_{RM}(x \to y) = \frac{s}{4} > 0 = c_{RM}(y \to x)$. Let $z = x$ and let $w = (0,t,t,0,0,0,...)$. Then, $\{z,w\} \in \mathcal{X}$ and $c_{RM}(z \to w) = \frac{s}{4} < \frac{3s}{4} = c_{RM}(w \to z)$.

Nearly Equal Interests: Let $s \in \mathbb{R}_{++}$, $s^+ = \frac{5s}{4}$, and $s^- = \frac{3s}{4}$. Let $\{x,y\} \in \mathcal{X}$ be such that $\{s(x,y,i) \mid i \in \mathbb{N}\} \cup \{s(y,x,i) \mid i \in \mathbb{N}\} \subseteq \{0\} \cup [s^-, s^+]$. Suppose $|\{i : s(x,y,i) > 0\}| > |\{i : s(y,x,i) > 0\}|$. Then, for any matching μ, there exists some i such that $s(x,y,i) \geq \frac{3s}{4} \wedge s(y,x,\mu(i)) = 0$ and thus $c_{RM}(x \to y) \geq \frac{3s}{4}$. On any rank-matching $\rho \in RM(y,x)$ and for any individual j, $s(y,x,j) > 0$ implies $s(x,y,\rho(j)) \geq \frac{3s}{4}$ and thus $c_{RM}(y \to x) \leq \frac{s}{2}$ as desired.

Nearly Innumerate: Let $s,t \in \mathbb{R}_{++}$ and let $n \geq 2$. Suppose:

$$\forall \{x,y\} \in \mathcal{X} : s(x,y) = s \wedge s(y,x) = (t,t) \implies x \succsim y \tag{1}$$

Let $x = (s,0,0,0,...)$ and let $y = (0,t,t,0,0,0,...)$. Then, $\max\{s-t,0\} = c_{RM}(x \to y) \geq c_{RM}(y \to x) = t$. Thus, $s - t \geq t > 0$. Then, for any $\{z,w\} \in \mathcal{X}$ such that $s(z,w) = s \wedge s(w,z) = \underbrace{(t,t,...,t)}_{n}$, $c_{RM}(z \to w) = s - t \geq t = c_{RM}(w \to z)$. Thus:

$$\forall \{z,w\} \in \mathcal{X} : s(z,w) = s \wedge s(w,z) = \underbrace{(t,t,...,t)}_{n} \implies z \succsim w \tag{2}$$

So (1) implies (2), and an exactly similar argument shows that (2) implies (1). For $\succ$, replace all the weak inequalities with strict ones. $\square$

A.2 Weakening the Assumptions

We now show that Fact 1 and Fact 2 can be obtained under much weaker assumptions than Definition 2 and Definition 5 about how the strength of one's residual complaint depends on the strength of one's interests and the interests (if any) of one's counterpart.

Let O be the set of all outcomes. Let S be the set of all strengths of an individual's interests, where $a \geq b$ means that interest a is at least as strong as b. Let $s : O \times O \times \mathbb{N} \to S$, where $s(x, y, i)$ is the strength of i's interest in x rather than y obtaining. We assume there is some minimum strength of interest $0 \in S$ that applies when, say, someone is wholly unaffected by the choice:

Principle 8 (Zero-Point)**.** $\exists 0 \in S : \forall s \in S : 0 \leq s$.

And we continue to assume that one's choices affect at most finitely many people:

Principle 9 (Finitude)**.** $\forall X \in \mathcal{X} : |\{i : \exists x, y \in X : s(x, y, i) \neq 0\}| < \infty \wedge |X| < \infty$

Recall that $\mathcal{M}$ is the set of all injections from $\mathbb{N}$ to $\mathbb{N}$. Let C be the set of all strengths of an individual's residual complaint, where $c \trianglerighteq d$ means that c is at least as strong as d. Let $c : \mathcal{M} \times O \times O \times \mathbb{N} \to C$, where $c_\mu(x, y, i)$ is the strength of individual i's residual complaint against choosing y over x under matching μ. We also assume:

Principle 10 (Ordering)**.** $\geq$ and $\trianglerighteq$ are orderings (i.e. total preorders) on S and C, respectively.

Principle 11 (Monotonicity)**.** If $s(x, y, i) \geq s(z, w, j)$ and $s(y, x, \mu(i)) \leq s(w, z, \rho(j))$, then $c_\mu(x, y, i) \trianglerighteq c_\rho(z, w, j)$.

With these principles, we can show (with essentially the same argument as before):

Fact 4. Zero-Point, Finitude, Ordering, and Monotonicity imply that for any $\{x, y\} \in \mathcal{X}$, $\mu \in RM(x, y)$ and $\rho \in \mathcal{M}$:

$$\max_{i \in \mathbb{N}} c_\mu(i : x \to y) \trianglelefteq \max_{i \in \mathbb{N}} c_\rho(i : x \to y)$$

Proof. Let $\{x, y\} \in \mathcal{X}$, $\mu \in RM(x, y)$, and $\rho \in \mathcal{M}$. Let $c^*_\mu = \max_{i \in \mathbb{N}} c_\mu(i : x \to y)$ and $c^*_\rho = \max_{i \in \mathbb{N}} c_\rho(i : x \to y)$. By Finitude, Ordering, and Monotonicity, these are well-defined.

Let $a = s(x, y, i^*)$ and $b = s(y, x, \mu(i^*))$. Since μ is a rank-matching, for any k such that $s(y, x, k) > b$ there exists a unique j such that $\mu(j) = k$ and $s(x, y, j) \geq a$. Thus,

$$|\{i : s(x, y, i) \geq a\}| > |\{i : s(y, x, i) > b\}|$$

Since ρ is injective, it cannot map the former into the latter. So there exists some $k^* \in \{i : s(x, y, i) \geq a\}$ such that $\rho(k^*) \notin \{i : s(y, x, i) > b\}$. By Monotonicity, $c_\mu \trianglelefteq c_\rho(x, y, k^*) \trianglelefteq c^*_\rho$. □

The proofs of Corollary 7 and Fact 1 proceed exactly as before.

These assumptions do not suffice for Fact 3 because we need stronger assumptions about the domain of possible choices $\mathcal{X}$ to vindicate several implications. Zero-Point, Finitude, Ordering, and Monotonicity are all consistent, for example, with the possibility that $S = \{0, 1\}$. But if $S = \{0, 1\}$, then Disparate Interests and Equal Interests are jointly inconsistent.[38]

References

Dorsey, Dale (2009). "Headaches, Lives and Value". In: *Utilitas* 21.1, p. 36.

Fleurbaey, Marc and Alex Voorhoeve (2013). "Decide as you would with full information". In: *Inequalities in health: Concepts, measures, and ethics* 2013, pp. 113–128.

Frick, Johann (2013). "Uncertainty and Justifiability to Each Person."" In: *Inequalities in health: Concepts, measures, and ethics* 129.

Gustafsson, Johan E. (2022). *Money-Pump Arguments*. Cambridge University Press. ISBN: 9781108718950.

Hammond, Peter J. (1988). "Consequentialist Foundations for Expected Utility". In: *Theory and Decision* 25.1, pp. 25–78.

Henning, Tim (2023). "Numbers without aggregation". In: *Noûs* 3, pp. 755–777.

Horton, Joe (2021). "Partial aggregation in ethics". In: *Philosophy Compass* 16.3, pp. 1–12.

Kamm, Frances Myrna (1984). "Equal treatment and equal chances". In: *Philosophy and Public Affairs* 14.2, pp. 177–194.

— (1993). *Morality, Mortality Volume I*. Oxford University Press.

— (2006). *Intricate Ethics*. Oxford University Press.

Korsgaard, Christine M. (2013). "The Relational Nature of the Good". In: *Oxford Studies in Metaethics*. Ed. by Russ Shafer-Landau. Vol. 8. Oxford: Oxford University Press, pp. 1–26.

[38] Strictly speaking, we consider analogues of these principles that replace $\mathbb{R}_{++}$ with $S \backslash \{0\}$.

Krantz, David et al., eds. (1971). *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York Academic Press.

Kumar, Rahul (2001). "Contractualism on saving the many". In: *Analysis* 61.2, pp. 165–170.

— (2011). *Contractualism on the Shoal of Aggregation*. Ed. by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon. . Oxford University Press.

Lazar, Seth (2018). "Limited Aggregation and Risk". In: *Philosophy and Public Affairs* 46.2, pp. 117–159.

Liao, S. Matthew and James Edgar Lim (2024). "Lives, Limbs, and Liver Spots: The Threshold Approach to Limited Aggregation". In: *Utilitas* 36.2, pp. 148–167.

McClennen, Edward F. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press. ISBN: 9780521360470.

Nebel, Jacob M. (2023). "The Sum of Well-Being". In: *Mind* 132.528, pp. 1074–1104.

— (2024). "Ethics Without Numbers". In: *Philosophy and Phenomenological Research* 108.2, pp. 289–319.

Otsuka, Michael (2015). "Risking Life and Limb: How to Discount Harms by their Improbability". In: *Identified versus Statistical Lives: An Interdisciplinary Perspective*. Ed. by I. Glenn Cohen, Norman Daniels, and Nir Eyal. Oxford University Press, pp. 77–93.

Pummer, Theron (2023). *The Rules of Rescue: Cost, Distance, and Effective Altruism*. New York: Oxford University Press.

Rüger, Korbinian (2020). "Aggregation with Constraints". In: *Utilitas* 32.4, pp. 454–471.

Scanlon, Thomas M. (1982). "Contractualism and Utilitarianism". In: *Utilitarianism and Beyond*. Ed. by Amartya Sen and Bernard Williams. Cambridge University Press, pp. 103–128.

— (1998). *What We Owe to Each Other*. Harvard University Press.

— (2021). *Contractualism and Justification*. Ed. by Markus Stepanians and Michael Frauchiger. De Gruyter. Reason, Justification, and Contractualism: Themes from Scanlon.

Steuwer, Bastian (2021). “Aggregation, Balancing, and Respect for the Claims of Individuals”. In: *Utilitas* 33.1, pp. 17–34.

Suppes and Winet (1955). “An Axiomatization of Utility Based on the Notion of Utility Differences”. In: *Management Science* 1.3-4, pp. 259–270.

Tadros, Victor (2019). *Localized Restricted Aggregation*. Ed. by David Sobel, Peter Vallentyne, and Steven Wall. Oxford Studies in Political Philosophy 5. Oxford University Press.

Taurek, John (1977). “Should the Numbers Count?” In: *Philosophy and Public Affairs* 6.4, pp. 293–316.

— (2021). *Reply to Parfit's 'Innumerate Ethics.' Principles and Persons. The Legacy of Derek Parfit.* Ed. by Jeff McMahan et al. Oxford University Press.

Temkin, Larry S. (2005). “A "New" Principle of Aggregation”. In: *Philosophical Issues* 15.1, pp. 218–234.

Timmerman, Travis (2015). “Sometimes There Is Nothing Wrong with Letting a Child Drown”. In: *Analysis* 75.2, pp. 204–212.

Van Gils, Aart and Patrick Tomlin (2020). *Relevance Rides Again?* Ed. by David Sobel, Peter Vallentyne, and Steven Wall. Oxford Studies in Political Philosophy 6. Oxford University Press.

Voorhoeve, Alex (2014). “How Should We Aggregate Competing Claims”. In: *Ethics* 125.1, pp. 64–87.

Weymark, John A. (1991). “A reconsideration of the Harsanyi–Sen debate on utilitarianism”. In: *Interpersonal Comparisons of Well-Being*. Cambridge University Press, pp. 255–320.

Zhang, Erik (2024). “Individualist Theories and Interpersonal Aggregation”. In: *Ethics* 134.4, pp. 479–511.