What If I Cannot Make a Difference (and Know It)*

Felix Pinkert

When several agents together produce suboptimal outcomes, yet no individual could have made a difference for the better, Act Consequentialism counterintuitively judges that all involved agents act rightly. I address this problem by supplementing Act Consequentialism with a requirement of modal robustness: Agents not only ought to produce best consequences in the actual world, but they also ought to be such that they would act optimally in certain counterfactual scenarios. I interpret this Modally Robust Act Consequentialism as Act Consequentialism plus a requirement of moral virtue, namely, to reliably act rightly and to act rightly for the right reasons.

We often collectively bring about bad outcomes.[1] For example, by continuing to buy cheap supermarket meat, many people together sustain factory farming, and the greenhouse gas emissions of millions of individuals together bring about anthropogenic climate change. Intuitively, these bad outcomes are not just a matter of bad luck, but the result of some sort of moral shortcoming. Yet in many of these situations, none of the individual agents could have made any difference for the better.

*I am grateful for the helpful comments on earlier versions of this essay from audiences at the Universities of Amsterdam, Munich, Oxford, Princeton, Stirling, and St. Andrew's. In particular, I am indebted to Ralf Bader, Marcia Baron, John Broome, Derek Ball, Mark Budolfson, Krister Bykvist, Cecile Fabre, Hilary Greaves, John Horty, Christian List, Brian McElwee, Paul McNamara, Tim Mulgan, Philip Pettit, Peter Singer, John Skorupski, Kai Spiekermann, Martin van Hees, Bill Wringe, and Donald Regan, another referee, and the editors of *Ethics* for illuminating comments on this topic.

1. When I speak of agents "collectively" or "together" bringing about an outcome, I use these terms in a minimal sense which only means that the individual actions of those agents happen to combine to bring about the outcome. I do not assume any sense of collective or deliberately cooperative action. Likewise, when I speak of collections and groups of agents, I use to term to include mere agglomerations.

Ethics 125 (July 2015): 971–998

Supermarkets only react to larger changes in demand than individual customers can effect, and someone choosing to take the train instead of the plane may not make any morally relevant difference to the climate if a given threshold of greenhouse gas concentrations will be, or has already been, exceeded anyway. Consequentialism, at least of the simple Act Consequentialist type, here seems unable to find fault with anyone, as no one could have brought about better results by acting differently. Consequentialists hence face the challenge of somehow still finding fault with agents for (morally) bad results which they together bring about.[2]

In a recent paper, Shelly Kagan takes up Derek Parfit's discussion of such "no-difference cases." Kagan argues that standard Act Consequentialism can meet this challenge when agents do not know if their contribution will make a difference.[3] He further contends that when agents do know that they cannot make any difference for the better, Act Consequentialists should concede that everyone acts rightly even though all agents together could have produced better outcomes.[4] Julia Nefsky, who strongly criticizes most other parts of Kagan's argument in her follow-up paper, does not take issue with this claim and likewise seems to assume that such cases of full knowledge pose no problem to Act Consequentialism.[5]

In this essay, I present a working example of agents who together bring about a collectively suboptimal outcome but who each know that due to each other's actions and dispositions, they individually cannot make any difference for the better (Sec. I). I then argue that contrary to Kagan's explicit and Nefsky's implicit position, we must find something morally objectionable about some of the agents (Sec. II). I argue that this challenge cannot be answered by Kagan's response, nor by any other common form of Consequentialism (Sec. III), and set aside an unnecessarily complicated and controversial solution (Sec. IV). Building on Donald Regan's and Michael Zimmerman's work, I propose further that

2. Henceforth, the qualification "morally" is implicit in all claims about the value of outcomes.

3. Shelly Kagan, "Do I Make a Difference?," *Philosophy and Public Affairs* 39 (2011): 105–41; Derek Parfit, *Reasons and Persons* (Oxford: Clarendon, 1984), chap. 3. Thresholds are only one of several possible reasons why individuals may be unable to make a difference in situations with multiple agents. The most important other such possible reasons are the seeming insignificance of small individual contributions (e.g., in pollution cases), and other agents' active frustration of one's attempts to improve outcomes (e.g., if other sellers will sell the weapons which one does not sell). While this essay does not discuss cases of seemingly insignificant contributions, frustration cases are covered by my proposed solution (see Secs. III.*A* and VII).

4. Kagan, "Do I Make a Difference?," 128.

5. Julia Nefsky, "Consequentialism and the Problem of Collective Harm: A Reply to Kagan," *Philosophy and Public Affairs* 39 (2012): 364–95.

these cases demand a radically different Consequentialist approach, according to which agents are required not only to act optimally in the actual world but also to be agents who would act optimally were others to act differently (Sec. V). I then show how this "Modally Robust Act Consequentialism" can be motivated by understanding it as Act Consequentialism plus a requirement of moral virtue (Sec. VI) and subsequently provide an inductive proof to show that it is a satisfactory response for a wide range of cases with any number of agents (Sec. VII). Finally, I discuss the limitations of my approach in cases of synchronous choice and contrast it with Donald Regan's Co-operative Utilitarianism (Sec. VIII).

I. THE TWO FACTORIES

Ann and Ben are owners of two factories which are located opposite each other on a river. Both agents opt for a production process which releases waste chemicals into the river and thereby kill all the fish in the river and destroy the livelihood of a fishing community downstream. The waste from one factory alone would suffice to kill all the fish, and adding the waste from the other factory does no additional damage whatsoever (say, the river flows into the sea where the waste is diluted below any harmful concentration). If Ann or Ben were to unilaterally produce cleanly, this would make their production uncompetitive compared to the other factory, put them out of business, and destroy the livelihood of their employees. However, if they both were to produce cleanly, then this problem would not arise, and both factories would remain in business and the fishing community would flourish. Ann and Ben each employ 100 workers, the fishing community counts 100 people, and all that matters morally in this case are the livelihoods of the workers and fishermen. Further, the only available actions are either to pollute or to produce cleanly. In particular, Ann and Ben cannot come together and suggest and discuss a common strategy.[6] So far, the case is represented by figure 1.[7]

6. I here set this possibility aside because the structure of the problem case can be reproduced on the level of discussing a strategy: If we assume for each agent that the other would not cooperate in trying to find a common solution, and that trying to start a discussion has morally relevant cost, then the payoff structure precisely matches that of the actions of polluting or producing cleanly. See Sec. VIII.*B* for a discussion of situations where agents can communicate.

7. Note that contrary to a typical strategic form game as used in game theory, we here only write one value per combination of strategies. This is because we are here only interested in how agents affect overall value, agent-neutrally considered. Thus the relevant "payoff" is the same for both agents.

		Ben	
		φ_c	φ_p
Ann	produce cleanly (φ_c)	all workers and fishermen have their livelihood (300)	only Ben's workers have their livelihood (100)
	pollute (φ_p)	only Ann's workers have their livelihood (100)	**only Ann and Ben's workers have their livelihood (200)**

Fig. 1.—The Two Factories as strategic form game

The two entrepreneurs find themselves in a suboptimal equilibrium in a coordination problem: In the two (Nash) equilibria where both agents pollute or produce cleanly, neither Ann nor Ben can improve the outcome by unilaterally acting differently. But from this fact alone, it does not yet follow that Act Consequentialism cannot find fault with Ann and Ben for their actions, because it is not yet clear that they could not have improved matters by acting differently simpliciter. For example, if Ann were someone who would produce cleanly if Ben produced cleanly, then Ben could improve matters by producing cleanly, and Act Consequentialism would judge that Ben acts wrongly by polluting. The Two Factories becomes a challenge for Act Consequentialism only once we assume that Ann and Ben are both "uncooperative," that is, each would pollute even if the other produced cleanly.[8] The additional stipulation of mutual uncooperativeness is needed because when Act Consequentialism evaluates the consequences of different actions available to an agent, it also takes into account how other agents would respond to these actions. In The Two Factories, it is only if both agents are uncooperative that neither could have improved matters by acting differently and that Act Consequentialism judges that both act rightly.

Lastly, Ann and Ben are fully aware of this situation. They know all the relevant facts that determine the consequences of the available actions, that is, the consequences of the different possible combinations of their actions and each other's actions and uncooperativeness. Hence they know that they individually cannot make a difference for the better by producing cleanly rather than polluting the river. These additional

8. For parallel examples, see Donald Regan, *Utilitarianism and Co-operation* (Oxford: Oxford University Press, 1980), 18; Christopher Woodard, *Reasons, Patterns, and Cooperation* (London: Routledge, 2008), 47; Michael J. Zimmermann, *The Concept of Moral Obligation* (Cambridge: Cambridge University Press, 1996), 257.

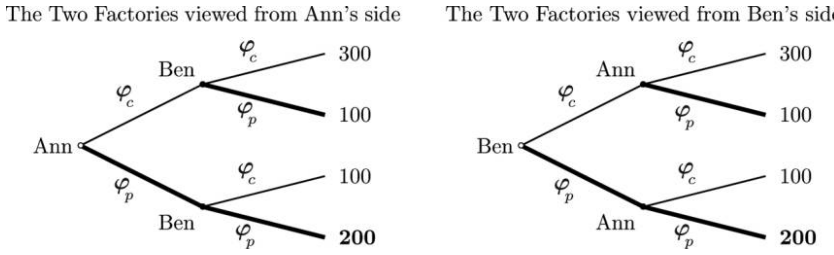

FIG. 2.—The Two Factories as extensive form game

assumptions of known mutual uncooperativeness are shown by the extensive form game in figure 2.[9]

II. CONSEQUENTIALISM AND COLLECTIVELY SUBOPTIMAL OUTCOMES

Cases like The Two Factories pose the following challenge to Act Consequentialism: Ann and Ben each individually could only have made matters worse by producing cleanly, as the other agent would then still have polluted the river, and the livelihoods of 100 workers in the cleanly producing factory would have been destroyed. Now Act Consequentialism is the following position:

> **Act Consequentialism:** Individual agents morally ought to perform those actions which have the best consequences.

Act Consequentialism judges that both Ann and Ben act rightly by polluting, even though they together could easily have brought about much better outcomes by both producing cleanly. So if Consequentialists only have Act Consequentialism to morally apprise a situation, then they let Ann and Ben off the hook for together producing collectively suboptimal outcomes. This is at odds with the following claim:

> **On-the-hook:** In any collection of agents who together gratuitously fail to bring about collectively optimal outcomes, there must be some relevant morally objectionable facts about some of the agents.[10]

9. As with the strategic form game, the "payoffs" represent overall value, agent-neutrally considered. Note further that the branches in the tree are meant to represent counterfactual relations and are not intended to represent any temporal order in which the agents act. However, as I discuss in Sec. VIII, certain morally desirable counterfactual relations presuppose that agents decide sequentially. I am indebted to Krister Bykvist for convincing me of the importance of tree representations.

10. In order to bypass the controversy about the good ought to be maximized or merely sufficiently much promoted, I stipulate that all discussed suboptimal outcomes are

By calling a failure to bring about optimal outcomes "gratuitous," I mean that the failure cannot be explained by mitigating circumstances due to which we could not expect a given group to collectively act optimally. Typically, such circumstances consist in non-culpable misinformation or lack of information. Since Ann and Ben know all relevant facts, I assume that their failure to bring about optimal outcomes is gratuitous.[11]

On-the-hook is a widely shared assumption in the philosophical discussion of Consequentialism and no-difference cases. Donald Regan's Prop-COP is the mirror image of On-the-hook, and Michael Zimmerman endorses the same claim, as does Derek Parfit in his discussion of collectively self-defeating moral principles. Shelly Kagan's and Julia Nefsky's recent discussions of no-difference cases are likewise motivated by an intuition along the lines of On-the-hook.[12]

On-the-hook is commonly defended along the following lines: The fundamental assumption of Consequentialism is that the right (i.e., the deontic status of actions) is solely determined by the good (i.e., by differences made to overall value in the world). In light of this connection, however, it should not be possible for each of us to act rightly while we together fail to produce optimal outcomes.[13] However, this motivation for On-the-hook is problematic, since it presupposes a Collective Consequentialist interpretation of the core Consequentialist position. According to this interpretation, the difference made that all agents can make together bears on the moral status of individuals' actions. Since this interpretation is denied by Act Consequentialists, basing On-the-hook on this interpretation already assumes that Act Consequentialism is mistaken. On pain of circularity, On-the-hook can then not pose a further challenge to Act Consequentialism.

I contend that for this reason, On-the-hook should not be understood as a specifically Consequentialist position. Instead, it should be understood as the contraposition of a second-order claim about morality in general and, hence, as a desideratum for any moral principle. According to this claim, the relation between morality and overall value is such that if everyone always satisfied all requirements posed on them by

---

so bad and so strongly suboptimal that even satisficing Consequentialists would concede that some fault needs to be found. Note further that "finding fault with" is always meant as a shorthand for "identifying morally objectionable facts about," and is not meant to carry any connotations of blameworthiness.

11. See the conclusion for cases with excusable collectively suboptimal outcomes.

12. Nefsky, "Consequentialism and the Problem of Collective Harm"; Kagan, "Do I Make a Difference?"; Regan, *Utilitarianism and Co-operation*, 4–5; Parfit, *Reasons and Persons*, 55–59; Zimmermann, *Moral Obligation*, chap. 9.

13. Parfit, *Reasons and Persons*, 54. See also Kagan, "Do I Make a Difference?," 108; Regan, *Utilitarianism and Co-operation*, 18–21; Woodard, *Reasons, Patterns, and Cooperation*, 45–47; and Zimmermann, *Moral Obligation*, 257–58.

morality, the world would be as good as it can be (as far as agents' influence is concerned). Whether or not these requirements are in any way determined by consequences, they "conspire" to direct all agents together toward the best outcomes they collectively can bring about.[14] Conversely, if collectively optimal outcomes are not produced, at least one agent must fall short of satisfying all the requirements of morality. When it is restricted to wrong action, as done by Kagan, Nefsky, and Parfit, On-the-hook captures the intuitive idea that it should not be possible for everyone to act rightly, yet for everyone together not to bring about optimal outcomes. I instead use a more general formulation of On-the-hook because, as I shall argue below, such an exclusive focus on actions as bearers of negative moral evaluation is misguided.

Thus understood, On-the-hook has considerable intuitive appeal, and a moral principle which can accommodate this intuition is, other things equal, strongly preferable to a moral principle that cannot accommodate it. So while I follow Kagan's claim that Act Consequentialists should see little appeal in condemning actions that are known to make the world no worse (see Sec. III.*B*), I hold that they should investigate other ways to accommodate On-the-hook. The extended form of Act Consequentialism which I propose is superior to Act Consequentialism and all other Consequentialist principles because it meets the challenge of accommodating On-the-hook without introducing any new problems. Before turning to my own proposal, however, I first discuss and reject two prominent nonsolutions and an unnecessarily complicated and controversial solution to this challenge.

III. NONSOLUTIONS

A. Subjective Act Consequentialism

Kagan offers a solution for those no-difference cases where agents do not know if their contribution will make a difference. Such situations typically arise when agents do not know the precise level of the threshold at which an additional contribution would make a difference, or they do not know how many other agents will contribute. According to Kagan, such cases call for a subjective Act Consequentialist assessment, where the deontic status of actions is determined not by actual but by expected consequences.[15] Now while this solution is effective for the cases Kagan

14. For the related "principle of moral harmony," which holds that moral obligations conspire to direct agents to collectively maximize a given group's welfare, see Fred Feldman, *Doing the Best We Can: An Essay in Informal Deontic Logic* (Dordrecht: Reidel, 1986), chap. 7, and "The Principle of Moral Harmony," *Journal of Philosophy* 77 (1980): 166–79.

15. Kagan, "Do I Make a Difference?," 125–27. Essentially the same argument is also advocated by Peter Singer, "Utilitarianism and Vegetarianism," *Philosophy and Public Affairs* 9 (1980): 325–37, 335–36.

considers, it does not apply to cases like The Two Factories, where the involved agents know that they cannot make a difference by acting differently. Here objective and subjective Act Consequentialism both judge that Ann and Ben each act rightly, and hence both moral principles fail to accommodate On-the-hook.

Kagan argues that this limitation of subjective Act Consequentialism is unproblematic because no-difference cases where agents know that they in fact cannot make a difference are rare and unrealistic. I contend that this claim is mistaken and that Kagan is misled by focusing on ordinary consumer choice and his unrealistic example of the Friends of Chicken Consumption—an organization which informs customers when their purchase of a chicken will not make any difference to the number of factory-farmed chicken in the world.[16] By contrast, realistic examples of no-difference cases with full knowledge are market situations where supply or demand of some good is known to be inelastic over a range of prices. Buyers and sellers then know that were they to stop participating, others would step in and compensate by taking over their transactions.

To give an example of price-inelastic demand, the number of weapons bought by a rich warlord and distributed to child soldiers is arguably determined largely by strategic considerations and the number of available children and is constant over a range of prices. Every individual weapons supplier can know that if they decide not to sell to the warlord, then another supplier will sell instead. With one supplier less in the market, the price of weapons may rise slightly, but this will not affect the total number of weapons bought nor the harm done to the child soldiers and to whoever gets into their way.[17] The same considerations hold for dealers selling highly addictive drugs. Examples of markets with price-inelastic supply are the purchasing of natural resources from conflict regions and (other) stolen goods. Unless Consequentialists are willing to bite the bullet and hold that there is nothing morally objectionable about agents on the price-elastic side of these collectively harmful and morally highly problematic transactions (i.e., those agents who cannot affect the amount of goods traded), they must meet the challenge of accommodating On-the-hook in no-difference cases with full knowledge.

B. Collective Consequentialism

One way to accommodate On-the-hook is to move away from Act Consequentialism to Collective Consequentialist moral principles. In determining what we individually ought to do, Collective Consequentialist

16. Kagan, "Do I Make a Difference?," 127–28.

17. The most likely response of the rich warlord to rising weapons prices is to reduce personal expenditure. This might make the warlord do less environmental damage by luxurious living, but this damage clearly is not the main reason for condemning such weapons sales.

principles consider the consequences of what we together can do. Examples of such principles are Consequentialist Generalization (which holds that agents ought to do what would have best consequences if everyone did it) and simple compliance-based Rule Consequentialism (which holds that agents ought to act in conformance with the set of rules which is such that if everyone complied with it, best consequences would ensue). These simple Collective Consequentialist principles accommodate On-the-hook by judging that Ann and Ben both act wrongly by polluting.

This achievement, however, comes at the price of judging that Ann and Ben each ought to unilaterally produce cleanly even though the other agent would then still pollute. Simple Collective Consequentialist principles then implausibly require Ann and Ben to perform actions which are known to produce no good and to destroy the livelihoods of 100 employees. I contend that a principle that accommodates On-the-hook by giving such reckless verdicts is not to be preferred over Act Consequentialism, which fails to accommodate On-the-hook but at least does not require agents to act recklessly.

These reckless verdicts can be avoided by moving to more sophisticated Collective Consequentialist principles, for example, by including caveats against recklessness or factoring in both the differences that we individually and that we together could make.[18] This response, however, boils down to modifying the principles to imply that Ann and Ben are not required to unilaterally produce cleanly and hence act rightly by polluting. Collective Consequentialists can thus avoid reckless verdicts only at the cost of not accommodating On-the-hook. The appropriate Consequentialist response to this dilemma then seems to be to opt for the lesser evil, namely, to not accommodate On-the-hook.

IV. A SOLUTION WITH BAGGAGE: COLLECTIVE WRONGNESS

The above dilemma occurs because so far, we have only considered Ann's and Ben's individual actions as potential bearers of moral fault, and these actions can only be either permitted or prohibited. Consequently, the only way to avoid the dilemma is to allow for objectionable facts about Ann and Ben which do not concern their individual actions of polluting.[19] A promising approach to extend the scope of moral evaluation is to let Act Consequentialism cover not only individual actions but also conjunctions of such actions across agents. According to this view, first proposed by Betsy Postow and Derek Parfit, it is Ann and Ben together who act

18. See, e.g., Tim Mulgan, *The Demands of Consequentialism* (Oxford: Clarendon, 2001), 60; Woodard, *Reasons, Patterns, and Cooperation*, 107–11.

19. This is the essence of the negative conclusion of Donald Regan's argument that no exclusively act-oriented moral principle can be fully adaptive (*Utilitarianism and Cooperation*, 105–23).

wrongly by both polluting, because together they could have brought about better outcomes.[20] Because asserting such collective wrongness is consistent with saying that Ann and Ben individually both act rightly by polluting, this position avoids the recklessness problem.

Whether this view also satisfactorily accommodates On-the-hook, however, is questionable. In its rough formulation so far, Ann and Ben individually are let off the hook, as there is nothing for which we could reproach either of them on their own. This problematic responsibility gap can be amended by stipulating that the collective wrongness reflects on Ann and Ben individually. On such a view, Ann and Ben act rightly qua individuals, but there is something wrong about them qua members of the group they constitute, due to the wrongdoing of that group.[21]

This version of the collective wrongness response, however, overshoots the target. Consider a variant of The Two Factories where Ben is uncooperative but Ann is cooperative. As far as Ann is concerned, the overall pollution is avoidable, as she would not pollute if Ben did not pollute either. However, Ben will pollute anyway, the fishermen still lose their livelihood, and Ann and Ben collectively act wrongly. It then seems implausible to hold that this fault reflects on Ann: After all, by being ready to bring about the collectively optimal outcome, she already does all that she can be expected to do. Furthermore, her fault qua member of the group of Ann and Ben together would arise solely from Ben's uncooperativeness. Ann's moral fault would then depend exclusively on facts about another agent, and Ann could not have avoided a moral deficiency even with her best efforts and intentions. This implication contradicts *ought implies can* and should thus be rejected.[22]

Consequently, for the collective wrongness response to be plausible, collective wrongness must reflect on individuals more selectively. The above discussion shows that we must discriminate between cooperative and uncooperative agents: Ben's uncooperativeness allows us to say that

20. Derek Parfit, "What We Together Do" (unpublished ms., 1988), 7; Betsy Carol Postow, "Generalized Act Utilitarianism," *Analysis* 37 (1977): 49–52.

21. See, e.g., Christian List and Philip Pettit, *Group Agency* (Oxford: Oxford University Press, 2011), 164; Marion Smiley, "Collective Responsibility," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, http://plato.stanford.edu/archives/fall2011/entries/collective-responsibility.

22. In such cases of unilateral uncooperativeness, Act Consequentialism can actually find fault with Ben, without any reference to collective wrongdoing: Because Ann is cooperative and would produce cleanly if Ben produced cleanly, Ben's action of polluting the river is suboptimal. So one could restrict the collective wrongdoing approach to account for cases of mutual uncooperativeness like The Two Factories where Act Consequentialism cannot find fault with either Ann or Ben, and could leave the remaining cases to be tackled by standard Act Consequentialism. However, such a patchwork approach seems implausibly gerrymandered. Furthermore, because the collective wrongdoing approach goes amiss in cases of unilateral uncooperativeness, the worry arises that even in cases of mutual

the collective wrongness reflects negatively on him, while the cooperative Ann must be let off the hook.[23]

In this improved form, the collective wrongness approach still relies on highly controversial ascriptions of collective wrongness and hence collective obligations. Even if we grant that some collectives of agents can be subject to moral obligations, it is highly doubtful that just any odd collection of agents can be subject to moral requirements. In order for the approach to fully accommodate On-the-hook, it would have to be shown that all cases of gratuitous failure to bring about collectively optimal outcomes are also cases where collective wrongness can correctly be attributed to the group in question.

I am doubtful that this challenge can be met. But rather than pressing this point further, in the following section I provide a solution that is superior either way. My proposal follows the improved collective wrongness approach by identifying agents' cooperativeness and uncooperativeness as facts that matter to the moral evaluation of these agents. However, my approach condemns agents' uncooperativeness directly, as opposed to viewing uncooperativeness as a mere conduit through which collective wrongness reflects on individuals. My approach thus bypasses the controversies that come along with referring to collective wrongness and has the further virtue of simplicity.

V. A BETTER SOLUTION: EVALUATING WHAT AGENTS WOULD HAVE DONE

Recall that it is Ben's uncooperativeness that makes Ann's act of pollution rightful, and vice versa. This is because by being agents who would pollute even if the other agent produced cleanly, Ann and Ben make it impossible for each other to achieve better outcomes by acting differently. A moral principle that condemns such uncooperativeness then finds fault in cases where Act Consequentialism cannot do so.[24] If Act Consequentialism is supplemented with such a principle, then it always

---

uncooperativeness where this approach gives the right verdicts, it does so for the wrong reasons.

The possibility of inescapable fault qua member of a collective that acts wrongly will not worry those who are at ease with "metaphysical guilt" that one bears solely in virtue of one's membership in a blameworthy collective (Smiley, "Collective Responsibility," sec. 4). This position, however, is rare, and at any rate not mine.

23. Note that the same qualifications would have to be added to a collective wrongness approach that does not evaluate Ann's and Ben's actions as a collective but other facts about them collectively.

24. In line with widening the scope of moral assessment beyond actions, I extend the use of the term "moral principle" to also cover general moral claims about what dispositions and characters agents ought to have.

finds fault when a group of agents fails to achieve collectively optimal outcomes—or so I will argue in the following.[25]

Now Ben's uncooperativeness means that even if Ann were to produce cleanly, Ben would not produce cleanly as well. In this counterfactual situation, Ben would act suboptimally, as he then could produce better outcomes by producing cleanly as well.[26] Hence while Ben satisfies the requirements of Act Consequentialism in the actual world, he would violate them in the counterfactual situation where Ann produces cleanly (and vice versa for Ann).

This connection between uncooperativeness and counterfactual violations of the requirements of Act Consequentialism holds generally: Abstracting from The Two Factories, call an agent "uncooperative" if she performs actions and has dispositions to act which make it impossible for others to bring about collectively optimal outcomes. For an agent to be uncooperative, it must hold that were all others to play their part in an optimal collective response, the uncooperative agent would, actively or passively, frustrate this attempt to reach a collectively optimal outcome. By doing so, the uncooperative agent would act suboptimally, because her actions would bring about a less good outcome than the collectively optimal outcome that would otherwise be produced. Hence the uncooperative agent would violate the requirements of Act Consequentialism.

Because of this connection between uncooperativeness and Act Consequentialism, we can capture the moral fault that is in play in The Two Factories by supplementing Act Consequentialism with a requirement of modal robustness, that is, a requirement to act according to the demands of Act Consequentialism not just in the actual world but also in certain counterfactual scenarios.[27] More precisely, the resulting requirement reads:

> **Modally Robust Act Consequentialism:** An agent ought to act optimally in the actual world, and be such that for all possible combinations of the actions of other agents, if that combination were instantiated, she would act optimally in these circumstances.

25. My proposed solution is in part inspired by Michael Zimmerman's "openness" requirement, which holds that agents ought to leave it up to others whether collectively optimal outcomes will be produced (Zimmermann, *Moral Obligation*, 263). Properly specified, Zimmerman's and my own views are extensionally equivalent, but I contend that my view is more straightforward, more easily motivated, and more suitable for generalization proofs. Also, my view fares better in situations of limited knowledge, as discussed in Sec. VIII.

26. As a terminological simplification, let the term "optimal" not only refer to best outcomes but also to actions or combinations of actions that among all relevant alternatives bring about optimal outcomes (and analogously for "suboptimal").

27. This formulation is adapted from Philip Pettit's 2011 Uheiro lectures on robustly demanding values or goods. Here, the idea is that values such as friendship do not just

Modally Robust Act Consequentialism accommodates On-the-hook in The Two Factories: Ann's uncooperativeness means that in counterfactual situations where Ben produces cleanly, Ann would act suboptimally, and likewise for Ben. Furthermore, this violation of the requirements of Modally Robust Act Consequentialism explains the suboptimal outcome, because universal satisfaction of the requirements of Modally Robust Act Consequentialism would have guaranteed collectively optimal outcomes, as follows: First, the modal robustness part of this principle requires Ann to be such that she would produce cleanly if Ben did, and vice versa. If both agents satisfy this requirement, then the Act Consequentialism part of the principle requires each agent to produce cleanly. If they also satisfy this further requirement, then they both produce cleanly, thereby producing collectively optimal outcomes. Conversely, if Ann and Ben do not produce collectively optimal outcomes, then this fact is explained by at least one of them violating the requirements of Modally Robust Act Consequentialism.

To strengthen the case for Modally Robust Act Consequentialism, in the following two sections I, first, independently motivate Modally Robust Act Consequentialism and, second, show that it accommodates On-the-hook not just in The Two Factories but in a wide range of cases. From here on, I use the adjective "cooperative" to refer to agents who satisfy the requirements of Modally Robust Act Consequentialism in a given situation and analogously for "uncooperative."

VI. MOTIVATING MODALLY ROBUST ACT CONSEQUENTIALISM

So far, I have supplemented Act Consequentialism with a modal robustness clause and have argued that the resulting Modally Robust Act Consequentialism accommodates On-the-hook in The Two Factories. In this section, I provide an independent motivation for Modally Robust Act Consequentialism. I argue that agents who satisfy the requirements of Modally Robust Act Consequentialism thereby display an important aspect of morally good or virtuous character, viewed from a Consequentialist perspective.

A. Consequentialist Virtue and Global Consequentialism

The position I advocate differs crucially from a common Consequentialist reading of moral virtue which evaluates dispositions and character

require certain behavior in the actual world but also in counterfactual scenarios. My proposal differs by applying the idea of modal robustness to moral requirements and hence moral rightness.

traits from a Global Consequentialist perspective.[28] On an individual-based form of Global Consequentialism, an agent's virtues are those dispositions and character traits which are optimal for her to have. As I show in the following, Modally Robust Act Consequentialism makes more stringent requirements on agents' dispositions than such a Global Consequentialism. Consequently, this kind of Global Consequentialist virtue does not capture the requirements of Modally Robust Act Consequentialism and, as I shall argue, is insufficient for fully accommodating On-the-hook in cases where Act Consequentialism on its own cannot find fault. I further show that if Global Consequentialism is modified to remedy this defect, the resulting position is implausible and hard to defend in its own right.

To begin with, note that by satisfying the requirements of Modally Robust Act Consequentialism, an individual agent does not always produce better outcomes than by displaying other dispositions. For example, in The Two Factories, outcomes would not be improved if Ann was cooperative and thereby satisfied the requirements of Modally Robust Act Consequentialism. This is because due to the fact that Ben uncooperatively pollutes the river, a cooperative Ann has to pollute as well if she is to save the livelihoods of her employees. The only difference between a cooperative and uncooperative Ann is that the former would make better outcomes available to Ben. But doing so would not make the world any better, as Ben would not use that new possibility. Hence, contrary to Modally Robust Act Consequentialism, individual-based Global Consequentialism here does not require Ann and Ben to be cooperative rather than uncooperative and also does not find fault with either of them.

The Global Consequentialist about virtue can respond by holding that dispositions need to be evaluated not in single choice situations but over longer time periods. Virtues are then those dispositions of an agent which in the long run lead to best outcomes. However, on this view, Modally Robust Act Consequentialism and individual-based Global Consequentialist virtue still come apart: First, if agents can have very fine-grained dispositions, then there are many competing possible dispositions that have just as good consequences as satisfying Modally Robust Act Consequentialism. For example, suppose that Ann and Ben are always cooperative except in The Two Factories. Because unilateral cooperativeness in The Two Factories does not make the world any better,

28. See, e.g., Julia Driver, *Uneasy Virtue* (Cambridge: Cambridge University Press, 2001), 67, 72; Parfit, *Reasons and Persons*, 24–26; and Philip Pettit and Michael Smith, "Global Consequentialism," in *Morality, Rules and Consequences: A Critical Reader*, ed. Brad Hooker, Elinor Mason, and Dale E. Miller (Edinburgh: Edinburgh University Press, 2000), 121–33, 121.

these dispositions have just as good consequences as fully satisfying the demands of Modally Robust Act Consequentialism. So individual-based Global Consequentialism cannot find fault with Ann and Ben for having these dispositions rather than satisfying the requirements of Modally Robust Act Consequentialism. It thus comes apart from Modally Robust Act Consequentialism and does not accommodate On-the-hook.

Second, even if we do not allow for extremely fine-grained dispositions, whether satisfying the demands of Modally Robust Act Consequentialism produces better outcomes in the long run than alternative dispositions depends on the prevalence of cooperativeness in other agents. If everyone were always disposed to frustrate any attempts to bring about better outcomes, then a wide range of choice situations would be similar to The Two Factories. Unilateral cooperativeness would then not bring about better outcomes than a range of alternative dispositions. Contrary to Modally Robust Act Consequentialism, individual-based Global Consequentialism would then not require agents to always be cooperative and, hence, cannot accommodate On-the-hook in such situations.

Lastly, Global Consequentialists can avoid these limitations by understanding virtues as those dispositions which it would be optimal for everyone to have. On this account, the requirements of Modally Robust Act Consequentialism are also requirements of virtue, because (as I will argue in the next section), universal satisfaction of the requirements of Modally Robust Act Consequentialism by all members of a collective guarantees collectively optimal outcomes. The downside of this view is that Modally Robust Act Consequentialism would then be motivated along Collective Global Consequentialist lines, which is an unattractive ad hoc patchwork of two theories that is hard to motivate: On this view, the deontic status of an agent's actions is solely determined by the consequences of that individual agent's actions, while the moral quality of her character traits is solely determined by the consequences of everyone having such traits. It is not clear why consequences should matter in such different ways when we assess actions and when we assess character.

B. Virtue and Responsiveness to Value

The alternative picture of the relation between Modally Robust Act Consequentialism and virtue starts from an independently plausible claim about the goodness of agents' characters. This claim holds that there is something better about agents who reliably act morally and are sensitive to those facts which are morally relevant than about agents who merely always perform the right action. This claim is independent of Act Consequentialism, and any moral principle can be supplemented by a modal robustness requirement: If a given moral principle is the right moral

principle, then there is something good about not just in fact satisfying the demands of that principle but in reliably doing so.

The value of such reliability is complemented by the value of the mental states that come along with it. Typically, agents reliably act rightly when they act for the right reasons, that is, when they "appropriately" and "effectively" take into consideration those facts which the moral principle identifies as morally relevant reasons that determine their moral obligations. Agents appropriately take into consideration the right reasons when they accurately assess their normative situation and care for the relevant reasons. They effectively take these reasons into consideration when this assessment translates into right action, as opposed to being hindered from doing so by weakness of will. When agents take the relevant reasons into account in this way, then their satisfaction of the demands of morality is not coincidental or morally lucky, and they would reliably satisfy these demands in various counterfactual scenarios. Hence modal robustness of acting rightly is arguably something good about an agent, due to the value both of reliably acting rightly and of the associated mental states that come along with such reliable right acting.[29] A requirement to modally robustly acting rightly can then be understood as a requirement of virtue: If an agent is to be virtuous, then they must display such modal robustness.

These claims about the goodness of agents' characters and the requirements of virtue holds no matter what specific moral principle is true: A moral principle like Act Consequentialism simply adds an account of which facts are the morally relevant reasons to which agents must be sensitive. The requirement of modal robustness is thus understood as a requirement additional to Consequentialist considerations. In particular, it is not a requirement of Consequentialism, that is, the claim is not that agents ought to display modal robustness because they thereby increase value (in the form of good character) in the world. Instead, considerations of virtue are an independent source of requirements on agents' characters.

Modally Robust Act Consequentialism understood in these terms has two attractive features. First, the moral defect manifested by uncooperative agents is distinct from morally wrong action: Just as a person with a psychopathic disorder might get through life without violating any deontological requirements, so uncooperative agents might be morally lucky and in fact always act optimally and according to Act Consequentialism rightly. Thus a requirement to be a morally good agent in this respect stands next to the initial Act Consequentialist moral principle that

29. For the value claim about acting for the right reasons, see Alison Hills, "Moral Testimony and Moral Epistemology," *Ethics* 120 (2009): 94–127, 108.

concerns right action.[30] It is this twofold moral assessment which enables Modally Robust Act Consequentialism to accommodate On-the-hook while avoiding reckless verdicts. Second, Modally Robust Act Consequentialism does not consider any more facts as morally relevant than Act Consequentialism does, as it holds that the only morally relevant reasons are the consequences of individuals' actions. Modally Robust Act Consequentialism interpreted as a requirement of moral virtue only adds the second-order requirement for agents to appropriately and effectively care about these very same reasons, and thus has the virtue of theoretical simplicity.

Modally Robust Act Consequentialism, motivated by appeal to moral virtue, then responds to the challenge of accommodating On-the-hook in The Two Factories as follows: Under mutual uncooperativeness, Ann and Ben each act rightly by polluting the river, because only in this way can they remain competitive and maintain the livelihoods of their workers. However, their uncooperativeness shows that there is something wrong with them as moral agents: They do not satisfy the demands of Act Consequentialism modally robustly, and this shows that they do not appropriately and effectively care about the livelihoods of the workers and fishermen. Because according to Act Consequentialism, the extent to which actions promote the good is the supreme (and only) moral consideration, and since a world with more people having livelihoods is better than a world with fewer people having livelihoods, Ann and Ben thus each individually show a morally problematic character trait.

C. Problems with Being Responsive to Value

The above connection between Modally Robust Act Consequentialism and moral virtue, however, is subject to a number of complications, which I can only outline here.[31] I have contrasted my approach with a Global Consequentialist approach and have argued that virtue is an independent source of requirements on agents' characters. However, even if one does not adopt a Global Consequentialist perspective, one can still assess character traits with regard to their instrumental value, for example, when one asks which character traits an agent ought to develop. Now it is commonplace that consciously caring for maximizing the good can have bad consequences. For example, it may disable one from living meaningful personal relationships or developing long-term commitments to projects whose value can only be realized if agents forgo constant reevaluation of their commitment. Conversely, dispositions other than caring for maximizing the good can have superior consequences. For example, aversion to killing can lead one to act suboptimally in some situations, but for

30. I am indebted to Brian McElwee for stressing this point.

31. I am grateful to Julia Driver for insisting on the importance of these complications.

most agents, such a disposition arguably has good consequences in the long run.[32] With regard to environmental coordination problems, Dale Jamieson has argued for countervailing considerations against counterfactual sensitivity toward others' actions. According to Jamieson, attempting to track others' actions and to calculate a precisely optimal course of action systematically leads to failure in large-scale coordination problems, due to moral hazard and negative psychological effects like cynicism.[33] In response to this problem, Jamieson argues, agents should reduce their individual contributions to environmental problems irrespective of what others do, even at the cost of not acting optimally in the actual world.[34]

Due to space limitations, I can only hint at an approach to this interplay between the requirements of virtue and the instrumental value of dispositions other than modally robust optimal action. I suggest that the concept of *pro tanto* requirements of virtue and higher-order caring for the good may be of help here. According this tentative view, on the "ground level" of agents' deliberation about actions, the requirement of virtue to appropriately and effectively care for the good is only a *pro tanto* requirement that can be outweighed by countervailing instrumental considerations against thinking in terms of maximizing the good. However, if Act Consequentialist is the correct account of the morally relevant reasons, then for agents to be fully virtuous, at some level of deliberation (e.g., when they deliberate about which character to develop), they must appropriately and effectively care for the good.[35] Otherwise, there would be a strange disconnect between the fundamental morally relevant reasons and morally virtuous agents.

In the following, I set these complications aside, and understand Modally Robust Act Consequentialism as Act Consequentialism plus an all-things-considered requirement of virtue to be a morally good agent by appropriately and effectively caring about maximizing the good. Just as people with psychopathic disorders who are insusceptible to typically deontological reasons are defective moral agents from a deontological standpoint, so uncooperative agents who do not appropriately and ef-

32. See Driver, *Uneasy Virtue*, 73.

33. Dale Jamieson, "When Utilitarians Should Be Virtue Theorists," *Utilitas* 19 (2007): 160–83, 167.

34. Jamieson's position is plausible because the negative effect of unilateral cooperative action is very small in the cases he discusses. For example, the moral cost of driving a smaller car even when everyone else keeps emitting too much carbon dioxide is very small, and is solely incurred by the cooperative agent. However, as The Two Factories shows, the cost of unilateral cooperativeness can be much higher, and in the case of institutional actors, the psychological effects discussed by Jamieson are likely to be less problematic. Moreover, it is hard to conceive of cases where additional considerations favor uncooperativeness, for example, to pollute no matter what others do.

35. There can still be virtuous agents who have the right ground-level dispositions but never thought about the overall good—only these agents are not as virtues as agents can be.

fectively care about promoting the good are then defective moral agents from a Consequentialist perspective.

VII. GENERALIZATION ABOUT MODALLY ROBUST ACT CONSEQUENTIALISM AND ON-THE-HOOK

As argued above, Modally Robust Act Consequentialism accommodates On-the-hook in the specific case of The Two Factories. However, there is a vast range of possible cases where several agents each have arbitrarily many options for acting and where they together gratuitously achieve only collectively suboptimal outcomes. In many of these cases, no individual could make any difference for the better by acting differently. If Modally Robust Act Consequentialism is to be a real improvement over Act Consequentialism in terms of accommodating On-the-hook, it must always find some moral fault in such no-difference cases. In this section, I prove that under certain epistemic assumptions, Modally Robust Act Consequentialism meets this requirement.

My proof proceeds by induction: I first show that Modally Robust Act Consequentialism always finds fault when two agents gratuitously bring about collectively suboptimal outcomes. Second, I show that if Modally Robust Act Consequentialism always finds fault in situations where some number n of agents gratuitously bring about collectively suboptimal outcomes, then it also does so in all situations with $n + 1$ agents. By induction, it then follows that the same claim holds in all situations with $n \geq 1$ agents.

A. Induction Start: Modally Robust Act Consequentialism in Situations with Two Agents

First, I establish the Induction Start, that is, the claim that Modally Robust Act Consequentialism always finds fault in any situation where two agents gratuitously achieve only collectively suboptimal outcomes, irrespective of the magnitude and distribution of outcomes, the number of optimal collective responses, and the number of actions that are available to each agent. So consider a case where two agents together gratuitously achieve only a collectively suboptimal outcome. Consider further all the different combinations of actions of the two agents which would have brought about an optimal outcome. Call all actions of a given agent which feature in at least one such optimal collective response her "potentially cooperative" actions. By contrast, "uncooperative actions" are actions which do not feature in any optimal collective response.

If there is only one optimal collective response, then each agent has only one potentially cooperative action available to her. Since only suboptimal outcomes are produced, it must hold that at least one of the agents performs an uncooperative action and must therefore have had a potentially cooperative action available to her which she does not per-

form. If there is more than one optimal collective response, then at least one of the agents has more than one potentially cooperative action available to her. In both cases, there is at least one agent who has had a potentially cooperative action available to her that she does not in fact perform, either because she performs an uncooperative action or because she performs another potentially cooperative action.

Now consider an agent a who could have performed such a (different) potentially cooperative action φ_c rather than the action φ_a that she actually performs, and ask whether she acted optimally by performing φ_a:

1. If she does not act optimally, then according to Act Consequentialism, she acts wrongly. A fortiori, Modally Robust Act Consequentialism finds fault with her.
2. If she does act optimally, then it must hold that had she performed the potentially cooperative action φ_c, outcomes would not have been better than the actual collectively suboptimal outcome. So had she performed φ_c, this would not have led to an optimal collective response. This is only possible if the other agent b would then not have performed a potentially cooperative action that would have completed an optimal collective response. Hence agent b would not have acted optimally in this counterfactual scenario. It follows that Modally Robust Act Consequentialism finds fault with agent b.

Hence either way, Modally Robust Act Consequentialism finds fault with at least one agent if a group of two agents gratuitously fails to produce collectively optimal outcomes.

B. Induction Step from n to n + 1 Agents

At this point of the argument, we need to make more precise the assumption that agents know all relevant facts, as follows:

> **Modally Robust Knowledge:** The involved agents know all relevant facts. If others acted differently, they would still have such knowledge, and this knowledge would likewise be modally robust.

The modal robustness requirement on agents' knowledge means that the agents cannot affect the quality of each others' epistemic positions, that is, deception or misleading each other are not among the available actions. This requirement needs to be recursive, that is, the modal robustness of an agent's epistemic position must itself be unaffected by others' actions, in order for step three of the below argument to be valid. Lastly, lest the assumption seems overly strong, for the purposes of my argument, "know" only needs to mean that agents have true beliefs and have such a high level of evidence that it does not make a difference to the verdicts of subjective

Act Consequentialism whether the true belief is treated as fact or only as true with a probability depending on the agent's evidence.

To establish the Induction Step, grant, for the sake of argument, the following Induction Assumption: If a group of n agents gratuitously produces collectively suboptimal outcomes, then Modally Robust Act Consequentialism finds fault with at least one of the group's members. Now consider a situation where $n + 1$ agents fall short of producing optimal outcomes. By the above considerations, there is then at least one agent a who has a potentially cooperative action φ_c available to her but who in fact performs another (potentially cooperative or uncooperative) action φ_a. Now we can again ask whether a acts optimally by performing φ_a rather than φ_c:

1. If a's performing φ_a is not optimal, then according to Act Consequentialism, a acts wrongly. A fortiori, Modally Robust Act Consequentialism finds fault with a.
2. If a's performing φ_a is optimal, then it must hold that had a performed φ_c, the remaining n agents would not together have completed an optimal collective response that includes a performing φ_c.
3. Now in this counterfactual situation, a doing φ_c is held fixed (as it is the antecedent of the above counterfactual). So we can reduce this counterfactual situation to a choice situation of the remaining n agents, while treating a doing φ_c as an external parameter. So had a done φ_c, the remaining n agents would have been in a situation where they together failed to bring about collectively optimal outcomes. Moreover, due to modally robust knowledge, in this situation, they could not have been excused due to ignorance of the actions of other agents, and their failure would thus have been gratuitous. Hence by the induction assumption, Modally Robust Act Consequentialism would have found fault with at least one of these remaining agents in the counterfactual situation where a does φ_c.
4. To see that Modally Robust Act Consequentialism also in fact finds fault with at least one of these agents, consider an agent b with whom Modally Robust Act Consequentialism would have found fault if a had done φ_c. Modally Robust Act Consequentialism would have found fault with b for one of two reasons. First, b would have acted suboptimally in the counterfactual situation where a does φ_c. In this case, b would have acted wrongly according to Act Consequentialism, and in the actual world, b violates the Modal Robustness requirement of Modally Robust Act Consequentialism. Second, b would have acted optimally but would have been such that in the counterfactual situation

where a does φ_c, it holds that there is then at least one further counterfactual combination of actions of the other $n-1$ agents to which b would not respond optimally. Modally Robust Act Consequentialism would then find fault with b on grounds of violating the Modal Robustness requirement. Now take one such counterfactual combination of actions, and add a doing φ_c to it. This gives us another possible combination of actions such that it holds in the actual world that were this combination instantiated, b would not respond optimally. Hence Modally Robust Act Consequentialism in fact finds fault with the would-be culprit.

5. So if a's performing φ_a is optimal, then Modally Robust Act Consequentialism finds fault with at least one of the remaining n agents.
6. Thus whether or not a's performing φ_a is optimal, Modally Robust Act Consequentialism finds fault with at least one agent.

The above argument gives us the Induction Step: If Modally Robust Act Consequentialism identifies some moral fault in all situations where n agents gratuitously fail to bring about optimal outcomes (the Induction Assumption), then it also finds fault in all situations where $n + 1$ agents gratuitously fail to produce collectively optimal outcomes. Now since Modally Robust Act Consequentialism finds fault in all situations where $n = 2$ agents gratuitously fail to produce collectively optimal outcomes (the Induction Start), by induction it also finds fault in all situations where any number $n \geq 2$ agents gratuitously brings about only collectively suboptimal outcomes. Hence given modally robust knowledge, Modally Robust Act Consequentialism accommodates On-the-hook in any such case with $n \geq 2$ agents.

Note that the above proof makes no reference to details of the strategic choice situation of agents, and hence does not assume that the situation is a coordination problem like The Two Factories. It thus applies to all cases where several agents together gratuitously bring about collectively suboptimal outcomes and have modally robust knowledge about each others' actions. These cases then include situations of active frustration, as in the market examples discussed in Sec. III.*A*.

VIII. LIMITATIONS AND RIVALS

Modally Robust Act Consequentialism is limited as a response to the challenge of accommodating On-the-hook because the above generalization proof relies on the assumption of modally robust knowledge. In this final section, I discuss cases where this assumption does not hold, and contrast Modally Robust Act Consequentialism with a prominent rival view which aims to do better.

A. Knowledge and Synchronous Choice

The assumption of modally robust knowledge breaks down in cases where agents cannot observe others' actual behavior when making their choices. Most notably, but by far not exclusively, agents face this situation in cases of synchronous choice, like the following specification of The Two Factories:

> **The Two Factories*:** Ann and Ben synchronously decide to pollute the river. When Ann makes her decision, she correctly infers from Ben's past conduct and expressions of intent that Ben will pollute the river.

Ann can here only base her decision on observation of Ben's past conduct and expressed intentions. But if, contrary to her best predictions, Ben were to produce cleanly, then there is no way for Ann to know about this fact when she makes her choice.[36] So all that Ann can do is to decide whether or not to pollute, depending on her beliefs about Ben's actions but independently from what Ben in fact does. She thus cannot both be such that she would act optimally were he to pollute and were he not to pollute. So in The Two Factories*, Modally Robust Act Consequentialism asks the impossible and hence cannot apply. Since modally robust knowledge can never hold in situations of synchronous choice, my discussion of Modally Robust Act Consequentialism so far has implicitly assumed that agents act in sequence.[37]

The standard way for Consequentialists to deal with situations with limited information is to link rightness to expected consequences. The Two Factories* shows that the same move is necessary for Modally Robust Act Consequentialism even in situations where agents in fact have full knowledge of others' actions but where that knowledge is not modally robust. Such Modally Robust Subjective Act Consequentialism requires Ann both to pollute in the actual world and to be such that she

36. I here assume that in the nearest possible world where Ben acts differently, Ann's epistemic position remains unchanged. I follow David Lewis by holding that this is possible even if it requires a "small miracle" ("Causation," *Journal of Philosophy* 70 [1973]: 556–67, 560).

37. Once we explicitly consider sequential choice, the following complication arises for my proof by induction: Those agents who get to choose earlier can determine not only the options faced by agents who get to choose later but might also influence the order in which others need to make their choices. In essence, my proof is unaffected by this possibility, as it makes no reference to the specific order in which agents choose, and does not assume that this order is fixed. Instead, the proof only implicitly assumes sequential choice due to its reliance on modally robust knowledge. This assumption is compatible with a variable ordering of agents' choices, and as long as this assumption is granted, the proof is unaffected.

However, cases with variable ordering of choices may well introduce significant cognitive challenges for agents to obtain modally robust knowledge. It may then be unrealistic

would pollute in the nearby possible world where Ben, contrary to her predictions, produces cleanly. This result is consistent with an interpretation of the principle via moral virtue: In The Two Factories*, the only way for Ann to be such that she would act optimally in the counterfactual situation where Ben produces cleanly is for her to produce cleanly in the actual world. But given that Ben in fact does pollute the river, this action of Ann's would be strongly suboptimal in the actual world. When morally good agents who care about the good are confronted with such a choice between either acting optimally in the actual world or acting optimally in counterfactual scenarios, they would act optimally in the actual world.[38] However, Modally Robust Subjective Act Consequentialism* cannot find any fault with Ann and Ben on the grounds that they would not act optimally in counterfactual situations where the other factory owner produces cleanly. Hence this view does not accommodate On-the-hook in The Two Factories*.

B. Co-operative Utilitarianism as Rival View

This limitation of Modally Robust Subjective Act Consequentialism makes Donald Regan's Co-operative Utilitarianism a particularly salient rival view, as it is explicitly developed for cases of synchronous choice. According to this view, agents ought to be cooperators who are willing to do their part in whatever the best response of all cooperators turns out to be. They further ought to follow a sophisticated procedure to identify the other cooperators and the best response available to all cooperators and to then play their part in this response.[39] Since the ambition of Co-operative Utilitarianism is to accommodate On-the-hook in synchronous choice cases, it promises to be superior to Modally Robust Subjective Act Consequentialism in these cases. However, as I argue in the following, this promise remains unfulfilled, as Co-operative Utilitarianism tacitly assumes a sort of sequential choice situation.

For Co-operative Utilitarianism to be applicable to cases like The Two Factories*, it must be possible for Ann and Ben to satisfy its requirements. Now identifying the other cooperators is not a purely mental activity that agents can do all on their own before making their decisions. Instead, it consists in approaching other agents with proposals for

to assume such beliefs, or, given the cost of obtaining such beliefs, the lack of such beliefs could be factor that makes agents' failure to produce optimal outcomes nongratuitous. These situations then provide an intermediate class of cases between situations where modally robust knowledge is plausible, and scenarios where it is impossible. I can here only express the conjecture that the future approach to cases of synchronous choice which I outline in the conclusion can be adapted to these intermediate cases as well. I am indebted to Donald Regan for pointing me to the possibility of variable orderings.

38. See Zimmermann, *Moral Obligation*, 266–68, for an analogous discussion with regard to Zimmerman's openness requirement.

39. Regan, *Utilitarianism and Co-operation*, chaps. 8–10, esp. 148 and 153.

collective strategies and in asking whether they would be willing to participate. So if we apply Co-operative Utilitarianism in The Two Factories* and other cases of synchronous choice, we implicitly assume that there is a prior choice situation in which agents can communicate their willingness and assess each others' willingness to cooperate. If such a prior choice situation is not given, Co-operative Utilitarianism must be modified to not require an impossible identification of other cooperators. And just like Modally Robust Subjective Act Consequentialism, such a modified Co-operative Utilitarianism* is no longer guaranteed to accommodate On-the-hook, because it can then easily be satisfied by agents like Ann and Ben who in fact act optimally in The Two Factories* but cannot engage in the process of identifying cooperators.

Furthermore, I contend that Co-operative Utilitarianism also does not have an edge over Modally Robust Subjective Act Consequentialism in situations of synchronous choice where prior communication is possible. This is because such prior communication allows agents not only to carry out the procedures required by Co-operative Utilitarianism but also to directly coordinate their actions.[40] Applied to The Two Factories*, agents then have a prior "coordination choice" to make, where they decide what to do with their ability to communicate, followed by the original synchronous "pollution choice," which concerns their actions with regard to the river.

Now if Ann and Ben first face a coordination choice, then Modally Robust Subjective Act Consequentialism can also accommodate On-the-hook: Suppose that Ann and Ben satisfy the requirements of Modally Robust Subjective Act Consequentialism in both choice situations. Now if in the pollution choice, they were confident that the other agent was going to produce cleanly, then they would both produce cleanly. Consequently, the best that Ann and Ben can do in the coordination choice is to create that confidence in each other by agreeing to produce cleanly. Since the coordination choice takes the form of a possible conversation, it allows agents to act in sequence. The problems of synchronous choice do then not arise, and we can assume modally robust knowledge of each others' actions. The subjective and objective forms of Modally Robust Act Consequentialism then coincide, and we can then refer to the above generalization proof and conclude that since Ann and Ben satisfy the requirements of Modally Robust Subjective Act Consequentialism, they will do the best they together can do in the coordination choice. Ann and Ben will hence agree to produce cleanly and then follow through with that agreement in the pollution choice and thereby produce collectively optimal outcomes. Conversely, if they do not bring about collectively optimal outcomes even though a coordination choice was avail-

40. Again I am indebted to Krister Bykvist for this observation.

able, then at least one of them violates the requirements of Modally Robust Subjective Act Consequentialism. Hence under the assumption that Ann and Ben face a coordination choice, Modally Robust Subjective Act Consequentialism and Co-operative Utilitarianism both accommodate On-the-hook.

My conjecture is that this result generalizes: Modally Robust Subjective Act Consequentialism and Co-operative Utilitarianism* can accommodate On-the-hook in the same cases, and so the limitations of the former do not give support to the latter. Given the potential for extensional equivalence of the two views, the main advantages of Modally Robust Subjective Act Consequentialism over Co-operative Utilitarianism are the conceptual simplicity of the former and the possibility of motivating it as a requirement of moral virtue.[41]

IX. CONCLUSION AND OUTLOOK

I have shown how Act Consequentialists can find fault with some agent in all cases where multiple agents who have modally robust knowledge of all the relevant facts gratuitously bring about collectively suboptimal outcomes, even if the agents individually cannot make any difference for the better due the uncooperativeness of others. Act Consequentialists, I have argued, can supplement their view of morally right action with a basic requirement of moral virtue, which holds that agents ought to be such that they act rightly not only in the actual world but also in counterfactual scenarios where others act differently. The resulting Modally Robust Act Consequentialism correctly finds fault with agents who fail to be morally virtuous in this respect and who thereby make collectively optimal outcomes inaccessible to others.

The main limitation of my proposal is its reliance on the assumption of modally robust knowledge, which breaks down in cases of synchronous choice. I have sketched how despite initial appearances, the competing view of Co-operative Utilitarianism fares no better than Modally Robust Act Consequentialism in such situations: I have conjectured that both views can accommodate On-the-hook in all and only those cases of synchronous choice that are preceded by coordination choices, but not in cases of isolated synchronous choice. A rigorous generalization of this claim, however, needs to wait for a future essay.

41. Regan argues that Co-operative Utilitarianism alone makes sense of Consequentialist morality as a communal enterprise of all cooperators together (*Utilitarianism and Co-operation*, chap. 12). Modally Robust Act Consequentialism, by contrast, remains firmly individualistic, and like Act Consequentialism treats the actions and dispositions of other agents in the same way as natural nonagential facts. I here have to leave it open if the insights of both positions can somehow be combined or subsumed under one of the theories.

Provided that such a generalization is possible, the main limitation of Modally Robust Act Consequentialism with regard to accommodating On-the-hook is that it does not apply to isolated synchronous choice cases where no prior coordination is possible. I contend that these cases can be accommodated by making independent use of the above motivation of Modally Robust Act Consequentialism via moral virtue. The idea is to approach these cases by means of epistemic game theory:[42] For any choice situation with a set payoff structure, this approach first identifies those sets of pairs of agents and probability assignments on others' actions for which universal satisfaction of Act Consequentialism is compatible with producing collectively suboptimal outcomes. For example, in The Two Factories* without a prior coordination choice, Act Consequentialism permits Ann and Ben to both pollute if they attach a probability of $\leq 1/3$ to the other agent producing cleanly. For each of these "epistemic profiles" of a possible situation, we then ask whether it is explained by a moral shortcoming of some agent which can be captured by the understanding of moral virtue advocated above (e.g., if Ben is known to be more likely than not to be uncooperative, because he does not care about the good, or if Ann negligently did not acquire enough information about Ben), or by mitigating circumstances which make collectively suboptimal outcomes non-gratuitous (e.g., if Ann mistakenly but non-culpably believes that Ben is uncooperative). If it can be shown that all problematic epistemic profiles can be explained by one of these factors, then Act Consequentialists can accommodate On-the-hook: The failure to bring about collectively optimal outcomes is then due to some moral shortcoming that Act Consequentialists can account for, or it is not gratuitous and no fault needs to be found. Developing this argument, however, is a task for yet another essay.

Finally, let me turn to the practical implications of my argument. As argued in Section III.*A*, the most realistic real-world scenarios where others' uncooperativeness is known to agents are market situations with price-inelastic supply or demand. Consider again the example of the rich warlord who will buy a set number of weapons for almost any price. What can Modally Robust Act Consequentialism say about the behavior and dispositions of a weapons dealer who considers offering weapons to the warlord? First, Modally Robust Act Consequentialism cannot say that the dealer would act wrongly by offering the weapons if the number of weapons offered by all dealers together is much greater than the warlord's demand. Second, the dealer ought to be such that were not enough others to offer their weapons, she would not offer them. At this point, we need a further, hithertho-not-discussed requirement that I can only stipulate here: a requirement to make sure that her cooperativeness is

42. I am indebted to Christian List for helping me specify the approach I am after.

known to the other dealers. This requirement can be satisfied in multiple ways. For example, the dealer may sign a contract requiring her to not offer her weapons if a given number of other dealers signs the same kind of contract. But the most simple way to make her willingness to not offer the weapons known is to simply not offer them in the actual world. Withdrawing her supply then is not an act intended to make the world better, but it is a communicative act by which the dealer signifies her willingness to exclude the warlord from any kind of weapons supply.

This application to a more concrete case is, of course, extremely sketchy. Further work is needed to justify a requirement not only to satisfy Modally Robust Act Consequentialism but to also make that fact known to others. Also, a requirement to make one's cooperativeness known can conflict with making the world better, for example, if one knows that other agents are not interested in cooperating anyway and if making one's cooperativeness known is costly (e.g., in terms of jobs lost in the dealer's company). The argument in this essay is hence only a very general starting point for detailed discussion of concrete cases. The important general lesson of my argument is that we need to evaluate not only agents' actual actions but also the actions they would perform if others acted differently and that, at least under favorable epistemic circumstances, such an extended moral evaluation captures all the moral facts needed to direct groups of agents toward collectively optimal outcomes.