How to Handle Trade-Offs in Pandemics

KRISTER BYKVIST
Stockholm University and Institute for Futures Studies

Abstract: Pandemics and other similar crises force us to make difficult moral trade-offs. It is tempting to think that this challenge should be met by invoking fundamental moral principles. This is a mistake. Instead, we need to work hard at designing institutions that enable the officeholders to make reasonable decisions under both fundamental ethical disagreement and empirical/evaluative uncertainty. It is argued that this is best done by supplementing the ethical-cum-legal platforms already in use with an ethical framework inspired by social welfare theory.

Keywords: trade-offs, pandemic, Covid-19, ethical disagreement, evaluative uncertainty, social choice framework, domain restrictions

JEL Classification: D63, D71, I38

I. INTRODUCTION
‘Follow the science’ has been a recurring catchphrase during the pandemic. It has been used by both politicians and scientists to justify radical measures to stop the spread of Covid-19. The obvious problem is that science on its own cannot justify measures—it can only tell you what is, will, or might be the case, and not what ought to be the case. It is tempting for us moral philosophers to think that we have exactly the right theoretical background to fill this gap between scientific results and the adoption of measures or strategies. After all, many of us work full time at constructing well thought-out and philosophically sophisticated fundamental ethical principles that will survive the criticisms of our peers in heated seminar discussions. It is therefore easy to think that the main focus should be on finding the most reasonable, universal fundamental principles. Once these have been identified and made public, the decision-makers can then use them to justifiably move from scientific results to decisions in pandemics and other crises.

AUTHOR’S NOTE: For very helpful comments on earlier drafts of this paper, I would like to thank Andrea Asker Svedberg, Orri Stefansson, Tim Campbell, and four anonymous reviewers for this journal.
I shall argue that this ‘filling the gap with a fundamental principle’ approach to pandemics and other similar global crises does not work, especially if the goal is to make us better prepared to deal with the next pandemic or global crisis. Rather, what we need to do is to work hard at designing institutions that enable the officeholders to make reasonable decisions under both fundamental ethical disagreement and empirical/evaluative uncertainty. This work must be a joint endeavour between moral philosophers, legal experts, politicians, and officeholders in agencies, and, crucially, it needs to gain political legitimacy through democratic processes. There is work to do for moral philosophers, however, especially when it comes to suggesting revisions of the various legally binding ethical platforms that are already in place for healthcare interventions in ordinary circumstances. But here we face a dilemma: either the platform rules are general and vague enough to gain broad societal approval, but this will come at the cost of not providing any clear guidance when it comes to difficult trade-offs, or they do provide such guidance but at the cost of lacking broad societal approval. I shall argue that we can avoid the sharpest horns of this dilemma if we supplement the ethical-cum-legal platforms with an ethical framework inspired by social welfare theory. An ethical framework is a system of rules that are in compliance with ethically binding platforms but provide more fine-tuned and precise trade-off principles, often by providing interpretations of the general and vague principles set by the platforms.

II. EXAMPLES OF TRADE-OFFS
It is clear that the Covid-19 pandemic has forced us to make many difficult trade-offs. Strict measures, such as lockdowns with bans on social gatherings and unnecessary shopping, can save many lives if they are put in place at the right time; but this comes at the cost of severe restrictions of freedom for many people whose wellbeing will also be negatively affected, including deaths caused by economic recession. Indeed, the measures will include restrictions of fundamental democratic rights, such as the freedom of assembly and freedom of movement. Since many of the lives saved will be vulnerable, older people, this trade-off is to a large extent one between the wellbeing and freedom of the young and the lives of the very old. But it is too simplistic to think it is just a matter of the young versus the very old. If we decide to impose minor restrictions with very little social distancing (or impose major restrictions at the last minute), the spread of the virus will cause much illness among younger
adults, and this is likely to overwhelm the health care system. This will have effects for the whole society because normal health care will not function properly. Furthermore, since it is easier for wealthy people than it is for poor people to cushion the negative effects of lockdowns and the widespread circulation of the virus, the trade-offs will also be between the wealthy and the poor, especially if there are no targeted efforts to help the poor. Finally, when designing vaccination plans, agencies will have to decide on a priority list. Difficult choices have to be made about whom to prioritize: those who have greater medical needs, those who are more likely to spread the virus further, or those who are needed to take care of the ill in elderly homes and hospitals. To make matters even worse, there is often grave uncertainty about the outcomes of the measures we take. So, in many cases, it is more accurate to say that we need to trade the risk of harm to one group against the risk of harm to another.

III. HOW TO UNDERSTAND TRADE-OFF QUESTIONS
The questions of how a certain trade-off should be resolved and how certain risks should be managed can be understood in more than one way. Moral philosophers tend to interpret them as asking for the trade-offs and risk managements that are morally right, and then they quickly turn to high-level moral theory for answers. More specifically, they tend to look for answers in some of the alternative foundational moral theories; for example, utilitarianism (maximize overall wellbeing!), Kantian ethics (act on maxims that you could will that everyone acts on!), and virtue ethics (do what the virtuous person would do!). There is nothing wrong with this, of course. High-level theorizing about the foundations of morality is important for the progress of moral philosophy, and it can also, if suitably popularized, provide important contributions to the public debate.¹ But the questions need not be read as asking for the right trade-offs and risk managements. Instead, they can be understood as asking for how agencies, governments, and other institutions should be set up to identify and resolve trade-offs and manage risks. This is still a moral question but it is one about institutional design and it should not be conflated with the former, more abstract ethical questions. Now, the ‘filling the gap with a fundamental principle’ approach need not conflate these questions, but it nevertheless assumes that once we have found a philosophically sound,

¹ See, for example, Public Ethics, a blog published by the Stockholm Centre for the Ethics of War and Peace (https://www.publicethics.org/).
fundamental ethical principle, it should be wheeled into the institutions and guide the decisions made by the relevant officeholders.

IV. Why is the ‘Filling the Gap Approach’ Problematic? One obvious problem is that this approach often cannot even get off the ground since there is no agreement among philosophers on fundamental ethical issues. There is thus no unique set of principles that can be offered to the decision-makers. It will not comfort the decision-makers, who are in urgent need of ethical guidance, to be told: ‘Be right with you, as soon as we have resolved the disagreements between utilitarians and virtue ethicists’ (slightly modified from Arras 2020, section 3.1). Indeed, there are even deep disagreements within the different moral traditions (for example, act utilitarians disagree with rule utilitarians, Neo-Kantians with Kantians, and Aristotelian virtue ethicists with non-Aristotelian ones, and so on). Hence, there is simply no hope that the community of moral philosophers will come to agree on a set of fundamental ethical principles that could be then presented to decision-makers before the next pandemic or global crisis.

Another problem is that no set of fundamental ethical principles is likely to gain widespread societal approval. There are two main reasons for this. First, fully worked out fundamental principles tend to be quite complex or presuppose a thorough understanding of certain philosophical concepts (for example, ‘final value’, ‘maxims’, and ‘practical virtue’). Many ordinary decision-makers will therefore have difficulties grasping and distinguishing between these principles (Kymlicka 1993, 4).

Second, what characterizes many fundamental principles is that they take a stand on difficult moral trade-offs and provide a specific explanation of why this is the right one to make. But this is exactly why any such principle will be deeply controversial in a pluralist society. There is no general agreement about difficult trade-offs, and, even when there is agreement, different explanations of why the trade-off is right are often provided. But if there is no widespread support for a fundamental principle, it will not gain political legitimacy through democratic processes.

Suppose we have somehow managed to jump these hurdles and have come to agree on a fundamental principle that has also gained societal approval and it is now to be presented to the decision-makers at an institution. Are we in the clear now? No, and one reason is that fundamental principles are difficult to apply to concrete cases (Kymlicka 1993, 6–8). To see this, take utilitarianism as an example. Suppose the decision-
makers want to apply it when deciding on a suppression or mitigation strategy for Covid-19. They then need to know whether the strategy maximises overall wellbeing. But how do they establish this? They need to know all the options they have and the consequences of each. In order to assess the consequences of an option in terms of overall wellbeing, they need to aggregate the individual wellbeing of all affected individuals by summing them up (or averaging them). It is no surprise that the decision-makers do not have all the relevant knowledge to make such a decision. It does not help much to go for a version of utilitarianism that tells them to maximize the expected overall wellbeing, for this still requires them to assess the aggregated overall wellbeing of all possible outcomes of each option. Furthermore, they need to decide on an account of wellbeing: Is it your pleasure that is good for you, or your preference satisfaction, or freedom, or your friendships, knowledge, and achievements, or something else? Here the decision-makers are likely to be undecided or in disagreement and thus face evaluative uncertainty or disagreement.

V. ETHICAL PLATFORMS

Instead of filling the gap between science and policy with a fundamental ethical principle, we can fill it with ethical-cum-legal platforms that have been approved through democratic processes. These platforms lay out in general terms certain ethical rules that have widespread support in society. These rules are supposed to be acceptable from very different ethical perspectives. From some perspectives, they capture (in part) what matters fundamentally, but from others they capture what is important instrumentally—that is, they are seen as rules we need to follow in order to respect or promote other more fundamental values. It is this ‘all-purpose’ feature of the rules that explains why they have gained wide societal support. Furthermore, since they are constructed with an eye to application in institutional settings (for example, hospitals), they have to be intelligible to and usable by the relevant officeholders and decision-makers.

Many countries have such platforms in place already, but they have not always been very useful during the pandemic. For example, the ethical platform for prioritizations in healthcare in Sweden is focused on individual-level medical interventions, such as giving one person priority for surgery over another (The Swedish Ministry of Health and Social Affairs 1996–1997). The overarching value is here the medical needs of the patient. One patient can be given priority over another only if the former has a greater medical need. It is not at all straightforward, however, to
apply this platform to measures in a pandemic, since many of these measures are applied at the population level, not the individual level, and it is not clear that medical need can always be the primary deciding factor. For example, in setting up a vaccination campaign for Covid-19, one may want to give extra priority to both vulnerable, older people with urgent medical needs and to certain members of the hospital staff, even though it is difficult to say that the latter group has urgent medical needs. The reason one may want to give priority to the hospital staff is not their urgent medical needs but rather the crucial role they play in supporting the healthcare system.

It is clear then that there is room for improving existing ethical platforms so that they are better poised to handle the kinds of trade-offs that a pandemic brings to the fore. Good proposals are found in SMER (2020) and Munthe, Heilinger, and Wild (2021). These proposals do not only provide a set of rules applicable to the specific problems facing us in pandemics, but they also identify the uncertainties and conceptual distinctions that are crucial to take into account when managing a pandemic crisis.

However, revising the existing platforms in light of these proposals would not avoid the dilemma presented in the introduction: they would be general and vague enough to gain broad societal approval but this would come at the cost of not providing any clear guidance when it comes to difficult trade-offs. For example, in Munthe, Heilinger, and Wild (2021, 4), we are told that “what to consider in that balancing [that is, trade-off] must be supported by a valid ethical principle”, but nothing is said about what that valid principle would be. Similarly, in SMER (2020, 11) we are told that we need to “minimize harm and save lives”, but nothing is said about how we should weigh harm against lives. How can we give the decision-makers more guidance?

VI. ETHICAL FRAMEWORKS
One way forward is to let the ethical platforms remain general and quite vague in their recommendations, since they need widespread approval through democratic processes, and delegate the more fine-tuned and precise ethical guidance to ethical frameworks that can be adjusted to specific institutional settings. But what should such frameworks look like?

One popular idea is to rely on a cost-benefit framework to make the necessary trade-offs. Here all moral values are transformed into monetary gains and losses and the aggregation procedure is basically one of
balancing these gains and losses. This is done by asking people how much they are willing to pay for various values, such as freedom, health, and happiness. The problems with such an account in its unadorned form are obvious. It masks ethical disagreements about moral trade-offs by transforming them into an economic trade-off problem where it may look obvious that we should not accrue more financial losses than the financial benefits can compensate for. What is more, since rich people are more willing to pay for these values than poor, values of the rich will be given more weight than those of the poor, when in fact these values are as important (if not more) for the poor. Finally, since all individual costs and benefits are supposed to be summed up, the framework does not take into account inequality and fairness, and so is far from uncontroversial.

A different framework looks at the aggregated loss of life years adjusted by their health quality (so-called QALYs). This framework seems more relevant for moral choice, especially if the quality-of-life years is not defined in terms of willingness to pay for a certain health status. But it will still be too insensitive to equality if the aggregation is simple summation. It will also give less weight to helping the worse-off, since the life years gained by people who would, if saved, be in good health generate a higher health-related quality-of-life than the life years gained by people who would live with disabilities or in poor health. One could also wonder how this framework is supposed to capture other important values such as freedom and achievements. Not all values that are at stake in pandemics can be reduced to health factors.

Common to these two frameworks is that they give very limited guidance on controversial trade-offs. Many moral factors are simply ruled out. This makes them ill-suited for decision-making where many different ethical perspectives must be taken into account. Of course, there is nothing wrong in using the frameworks as one of many inputs in deliberation, but they cannot be the only input.

Is there a better framework that is still action-guiding but more inclusive of different ethical views? Recently, Adler et al. (2020) have proposed a social welfare analysis that could be used as an ethical framework for moral trade-offs in pandemics.² Individual wellbeing is measured by a bundle of individual goods that can be compared across people, and moral choice is represented as a function of these bundles. More precisely, the input for this function are \( n \)-tuples of individual wellbeing values \((v_1, v_2, \ldots, v_n)\), where \(v_i\) is the numerical representation of the wellbeing

² Adler (2019) argues that this framework should be used for all policy-decisions.
of individual $i$. The output is a ranking of bundles that can then be used for moral choice—it is permissible to bring about a top-ranked bundle. (Alternatively, the ranking step is bypassed, and the output is simply a set of permissible options.) Traditionally, social welfare analysis has been focused on preference-based wellbeing, but it can be generalized to take into account other, more objective values such as freedom and achievements. Furthermore, the choice function need not be one of simple summation. One can, for example, give more weight to worse-off people, and one can represent the disvalue of inequality in different ways (either as an individual harm or as an impersonal disvalue of the distribution).

One problem with this framework is that it is quite formal and mathematical and therefore requires training to be used effectively. Decision-makers will not have time for this, and if they defer to experts, they will have a hard time understanding why a certain option comes out as best in the framework. Another problem is that we might still be far away from any clear guidance on how to resolve trade-offs. Even if it is possible to represent diverse views by listing different values for each person and construct different choice functions, this is not enough to provide any clear action-guidance. It may clarify what the disagreements are all about, but the decision-maker needs more guidance.

Things are not hopeless, however. The problem with the high level of formalism can be solved to some extent if we focus on less mathematical and more informal versions of the social choice framework, where diagrams and other pictorial methods can facilitate understanding. There is still need of guidance on this, of course, but that could be given by philosophers and social scientists serving on ethical boards within the relevant institutions. So far, I have been sticking closely to the framework and ideas presented in Adler et al. (2020). In the following, I will introduce some ideas that go beyond those discussed by them.

The problem of not providing any action-guidance can be lessened somewhat by using a methodology that is part and parcel of social choice theory, namely, to identify a set of conditions of adequacy for any acceptable moral trade-off principle. Examples of such conditions are Pareto conditions (for example, making everyone better off is an improvement), non-dictatorship (no single individual’s interests are decisive), independence (what is an acceptable trade-off between two individuals regarding the choice between two options should not be affected by how the individuals fare in other options), and acyclicity (betterness cannot be cyclical: it cannot be that $A$ is better than $B$, which is better than $C$, which is better than...
The benefits of focusing on adequacy conditions is that it is likely that there is more agreement on and less uncertainty about them, for they provide only minimal constraints on acceptable trade-offs and are silent on why a certain trade-off is acceptable or not. For example, both utilitarians and virtue ethicists can accept the Pareto condition, but for different reasons. These conditions can then be used to weed out certain principles that violate all conditions. Among the remaining principles, a partial ranking can be generated: if a principle $P_1$ satisfies all the conditions that another principle $P_2$ satisfies, but in addition satisfies some other conditions, $P_1$ is to be preferred to $P_2$. Of course, this will not generate a complete ranking, since one principle may do better under one set of conditions and another principle will do better under a different set of conditions.

This can be remedied somewhat if the decision-makers agree on a ranking of the conditions in terms of their importance, for—other things being equal—violating a more important condition is worse than violating a less important condition. Perhaps acyclicity is seen as more important than independence, for example, since violating it threatens the very possibility of consistent choice. Note also that legally binding ethical platforms may rule in favour of certain conditions of adequacy but rule out others, for instance, those that together imply some verdict that goes against the principles in the platform. The ones that are ruled out are thus not just less important, but they cannot be invoked at all.

We still cannot expect a complete ordering, but we may not always need one for action-guidance. If a principle is ranked lower than all others, it can simply be discarded. If two principles, which cannot be ranked against each other, are ranked higher than all others, following either may seem to be a sensible choice at least in some circumstances. When a unique decision is required, the decision-makers with ultimate accountability, for example, the government, will have to make a choice between the top-ranked ones.

Those who are familiar with the literature on social welfare will now demur: Is not this reliance on conditions of adequacy a non-starter since it is well-known that there is no principle that satisfies them all? After all, so-called impossibility theorems are ubiquitous in the social welfare tradition. Even if we allow for interpersonal comparisons of welfare gains and losses and thus avoid Arrow’s original theorem, we are not out of the

\footnote{For more on the role of conditions of adequacy in social welfare theory, see Gaertner (2009).}
woods. As Fleurbaey, Tungodden, and Vallentyne (2009) and Fleurbaey and Tungodden (2010) show, even with this rich wellbeing information, there is no moral trade-off principle that satisfies a set of very plausible conditions, including the very plausible requirement that a huge welfare loss to a few worse-off people cannot be compensated for by small welfare gains for many very well-off people. Similarly, Arrhenius (forthcoming) and Blackorby, Bossert, and Donaldson (2005) show impossibility theorems for principles that take into account the wellbeing of future generations and yet allow for rich welfare information.

So, it is true that one tends to face impossibility theorems when constructing universal trade-off principles, but for these theorems to work they need to assume a rich domain of bundles of individual goods. There are possibility results that show that by restricting the domain it becomes possible for a principle to satisfy all conditions. This is good news for the application of the social choice framework to institutional settings, since it is clear that the domain will be severely restricted in these contexts. First of all, there are practical constraints; the institutions are quite restricted in what they can or cannot do (to a large extent set by budgetary constraints), so some bundles of individual goods are simply not feasible. But there are also legal constraints. The government tells the institutions what they have mandate to do, and legally binding ethical platforms will set the legal limits for the trade-offs that can be considered. Some bundles will be ruled out because there is no legally acceptable way to bring them about. One example is trade-offs that would violate some citizens’ legal rights. This means that the domain that is left for the ethical framework is likely to lack the richness that is needed for the impossibility results.

Of course, this means that the universality of the trade-off principles no longer can be maintained. But this is exactly what we want. Recall that the role of these frameworks is to facilitate decisions in contexts where very different ethical perspectives are represented. The focus must therefore not be on finding universal principles, but on finding workable principles for limited domains. Indeed, moral philosophers in the bioethics traditions have already done some useful work here, searching for what

---

4 For more Arrovian impossibility theorems in social choice theory, see Gaertner (2009). Note that these impossibility theorems are relevant even if interpersonal wellbeing comparisons are meaningful, for the decision-maker might lack information about such comparisons.

5 See, for instance, Gaertner (2001) for possibility results when wellbeing cannot be compared across people. See Arrhenius, Budolfsson, and Spears (2021) for a possibility result when wellbeing can be compared across people.
they often call ‘mid-level principles’ (Arras 2020, section 6.4.1). A lot of this work is wedded to a sceptical attitude towards abstract theorizing in moral philosophy. But you do not have to be a theory-sceptic in order to be interested in limited domain questions. Indeed, working out principles for ethical frameworks with limited domains can be seen as a form of **modelling** of certain parts of the moral landscape akin to the scientific modelling of certain restricted natural phenomena, such as the behaviour of the fish population in a certain lake. Common to both modelling approaches is that: (a) they are local and do not claim to capture phenomena outside the limited domain, but at the same time they do not rule out the possibility of finding more fundamental explanations that have a larger scope; (b) many factors are ignored because they are irrelevant, too complex to handle, or unable to make a significant difference; (c) some formal representations are used, often mathematical in nature but they need not be. If this kind of modelling appeals to theory-friendly scientists, it should also appeal to theory-friendly moral philosophers.

So far, I have suggested ways that evaluative disagreement and uncertainty about trade-off principles can be handled in ethical frameworks. But we have still to tackle the problems of application. Say we have found a set of eligible trade-off principles, perhaps principles about how to weigh wellbeing against freedom. We are still unable to apply these principles since there are many different ways to understand wellbeing and freedom. Does wellbeing consist in pleasure, preference satisfaction, or happiness? Does freedom consist in having options, so that the more options you have, the more freedom you enjoy, or does your degree of freedom also depend on the value of your options? Again, there is no chance that moral philosophers will have anything like an agreed stance on these issues. Even if they did, their views would not gain broad societal approval. The ethical frameworks thus will have to be as neutral as possible on these fundamental matters.

One way they can be more neutral is by focusing on **proxies** for wellbeing and freedom, factors that to some significant degree co-vary probabilistically with the respective value. Reliance upon proxies is unavoidable, since no one can directly measure every single individual’s wellbeing and freedom in society. This restriction can be turned into an advantage, however, for one can choose proxies for wellbeing or freedom that are

---

6 This analogy between normative and scientific modelling is developed in Roussos (2020).
7 One of the current projects of the author and Joe Roussos is to develop this point.
sufficiently coarse-grained to co-vary with different candidates for well-being or freedom. For example, answers to self-reported happiness questionnaires co-vary to a significant degree with happiness seen as life-satisfaction, but also with pleasure and preference-satisfaction, since when you are overall satisfied with your life you tend to feel pleasure and have satisfied preferences. Similarly, some freedom measures, such as answers to self-reported freedom questionnaires, will not be able to clearly distinguish between different notions of freedom.

The final problem with applying trade-off principles is how to deal with empirical uncertainty. We often do not know how a certain pandemic measure will affect people's wellbeing and freedoms (or their proxies thereof). How should we weigh the risk of something bad against the chance of something good? Or, more specifically, when is it permissible to impose a certain risk of a freedom restriction for a certain group of people for the sake of wellbeing benefits for this or some other group of people? Here, again, we can expect disagreement. In fact, the questions of which risks are acceptable to impose on others are moral questions. So, what we have here is another source of a moral disagreement. But one thing people tend to agree on is that the persons the risks are imposed upon should have a say. This is one important aspect of how medical treatments and experiments are regulated in most countries. So, ethical frameworks need to be sensitive to this. Another thing to say is that since risk-imposition is to a large extent a moral issue, risk assessments should not be dealt with separately from the trade-off principles. Rather it is often better to formulate the trade-off questions directly in terms of risks: Is it permissible to impose a certain risk of harm to some people for the sake of certain expected gains for others? Disagreements and uncertainty about these risky trade-off principles can then be handled in a way that is similar to the non-risky trade-off principles, namely, by constructing a partial ordering of principles by invoking conditions of adequacy, ranked in terms of importance.

An important caveat: I do not pretend that this ethical framework will provide us with clear guidance for all controversial trade-offs. Disagreements and uncertainty, both about facts and trade-off principles, and even about some conditions of adequacy, will persist in many cases. That is why I can only claim to avoid the sharpest horns of the initial dilemma. The framework can only clarify and reduce these disagreements and uncertainties. We also need to rely on negotiations and compromises, but
they are unavoidable parts of any collective decision-making about contested issues.

VII. CONCLUDING REMARKS

Much more needs to be said before we have a detailed proposal of the kind of ethical framework that needs to supplement ethical platforms. For example, more needs to be said about the interplay between legally binding platforms and ethical frameworks, since the platforms to a large extent restrict the domains of the frameworks. But I hope that I have said enough to show the potential advantages of this approach for resolving trade-offs in pandemics and other crises.

REFERENCES


**Krister Bykvist** is a Professor of Practical Philosophy at Stockholm University and a Research Fellow at the Institute for Futures Studies, Stockholm. His research concerns questions about our responsibility for future generations, the foundations of consequentialism, evaluative uncertainty, and the relationship between preferences, value, and welfare. His most recent book is *Moral Uncertainty* (co-authored with Will MacAskill and Toby Ord), published in 2020 by Oxford University Press.
Contact e-mail: <krister.bykvist@philosophy.su.se>