

Following the Science: Pandemic Policy Making and Reasonable Worst-Case Scenarios

Richard Bradley

Joe Roussos

Abstract

The UK has been ‘following the science’ in response to COVID-19 pandemic in line with the national framework for the use of scientific advice in assessment of risk. We argue that the way in which it does so is unsatisfactory in two important respects. Firstly, pandemic policy making is not based on a comprehensive assessment of policy impacts. And secondly, the focus on reasonable worst-case scenarios as a way of managing uncertainty results in a loss of decision-relevant information and does not provide a coherent basis for policy making.

Key words: Scientific advice, policy making, reasonable worst-case scenario, uncertainty

JEL number: can be found at <http://www.aeaweb.org/jel/guide/jel.php>

Corresponding author: Richard Bradley, r.bradley@lse.ac.uk

1. Introduction

The COVID-19 pandemic provided policy makers with few precedents in deciding how to act. While influenza pandemics are not uncommon and while there have been other recent coronavirus epidemics (notably, SARS and MERS), the nature of COVID-19 meant policy makers had to act amid conditions of massive uncertainty; in particular, about the characteristics of the new virus, the medical and societal effects it would have and what kind of interventions would be effective against it.

In many countries, policy makers rightly turned to science, and the mantra that the governments were ‘following the science’ became ubiquitous. The UK’s initial response was guided by the Scientific Advisory Group for Emergencies (SAGE), a group consisting of experts (mainly epidemiologists, clinicians, virologists and behavioural scientists) drawn primarily from scientific-academic institutions, but also from public sector bodies, industry and commerce. At times, SAGE’s scientific advice has been decisive. For example, a report of the Imperial COVID-19 Response team [7] resulted in a significant change in policy direction, moving the government to pursue active suppression instead of controlled acquisition of herd immunity. At other times,

SAGE's advice has been less influential. The UK's late 2020 lockdowns, for instance, were implemented later than SAGE advised.

Political leaders' claims to be simply following scientific advice have perhaps been motivated, in part, by a desire to avoid responsibility for the outcomes of policy decisions, given the real possibility their decisions would turn out to be the wrong ones (something that is inevitable in situations of severe uncertainty). But following the science is not the same as letting scientists decide, a point on which scientific advisors in the UK have been clear and insistent. 'Scientists advise, politicians decide' has been a frequent counterpoint to the politicians' expressions of deference to science.

What then is the appropriate relationship between policy making and scientific advising? It is evident that scientific advice alone cannot settle the issue of what policy to adopt. For this political decision must be sensitive to judgements about how good or bad any possible policy outcome is, and it is simply not the role of science to provide such value judgements. Moreover, even given the requisite values, pandemic policy choices are not fixed by the kind of scientific advice that can be offered in circumstances of scientific uncertainty, such as those characterising the pandemic. In these circumstances scientists cannot (with any confidence) simply say what the outcome will be of any particular policy choice, or even assign precise probabilities to these outcomes. For this reason, SAGE largely eschews predictions and instead offers scenarios considered plausible. But being advised that something is a plausible outcome of a policy does not tell one whether to adopt the policy or not, even given full knowledge of how desirable the outcome is.

Our question is then: How should policy follow the science when the scientific advice fails to resolve uncertainty about the outcomes of policy choices? To answer it, a match must be found between the kind of advice that scientists can provide in circumstances of severe uncertainty and the kind of input that policy making requires. Such a match exists, for instance, when policy makers seek to maximise expected social wellbeing and scientists can provide analysis that shows how any given policy will likely affect key measures of wellbeing. Indeed, this is just the match prescribed by standard decision theory. But it is not one that is available in the context of the COVID-19 pandemic.

In this paper we look at the nature of the relationship in the UK between pandemic policy making and the form of scientific advice being offered. We will argue that the current match is not satisfactory for two reasons. Firstly, the scientific advice is not sufficiently integrated with respect to the different determinants of policy success. Public health outcomes have dominated discussion, where what is needed is an assessment of health outcomes together with economic and social wellbeing. This assessment should be integrated, in the sense that dependencies between these factors should be explicitly accounted for. The right policy is the one which delivers the best outcomes *all things considered*, rather than merely the best public health outcomes. Secondly, the form of scientific advice, and in particular the focus on a 'reasonable worst-case scenario' as a means of managing uncertainty, does not provide an adequate guide for policy decision making. It focuses too much attention on a single scenario and fails to provide decision

makers with crucial information. Although these lessons are drawn from the examination of the UK's response to the COVID-19 pandemic, we believe that they are of general relevance to debate about the reform of arrangements for the conscription of scientific advice for responses to emergencies.

2. Scientific Advice

Let's start with an idealised description of the problem of pandemic policy making in times of uncertainty, as faced in the UK. A policy decision maker (or decision-making body) must choose amongst a range of feasible and morally and legally permissible alternative policy responses to the pandemic. They seek to do so with primary regard to the impact of the pandemic on people's wellbeing. Let us assume that individual wellbeing depends on multiple factors, such as the mental and physical health of the individual, their lifespan and income, and potentially factors such as the opportunities and liberties they are afforded. These factors, and distributions of combinations them (what we will call comprehensive outcomes), are the ultimate targets of policy interventions.

To make an adequate policy choice, two questions must be addressed. The first is the impact of the pandemic, and the potential policy responses to it, on the population's wellbeing. This is an empirical question to be settled by scientific enquiry. For example, how many are expected to die absent intervention? Or what is the educational impact of closing schools for a month? The second question is what evaluative weight to give to the different comprehensive outcomes. This is primarily a normative question, although it has empirical aspects. For example, is such-and-such impact on education worse than some number of deaths? On both questions the decision maker can and should seek expert advice: the advice of health experts on the public health effects, of behavioural scientists on the responses of individuals to policy measures, of economists on the effect on incomes, and so on.

In the UK, the arrangements for acquiring scientific advice to support the government responses to emergencies are set out in the National Risk Assessment (a classified document summarised in the Risk Register [11]). It was under these terms of reference that SAGE was activated in January 2020. The role of SAGE was, and continues to be, to ensure that "timely and coordinated scientific advice is made available to decision makers to support UK cross-government decisions in the Cabinet Office Briefing Room (COBR)"ⁱ and for "coordinating and peer reviewing, as far as possible, scientific and technical advice to inform decision-making"ⁱⁱ in the face of emergencies.

SAGE draws on the work of specialist sub-groups. Public health modelling is done by the Scientific Pandemic Influenza Group on Modelling, Operational sub-group (SPI-M-O), while the modelling of the people's responses to policy interventions is done by the Scientific Pandemic Insights Group on Behaviours (SPI-B). SPI-B has been the primary, and perhaps sole, social scientific contributor to SAGE during the pandemic. Another advisory group with similar make-up, NERVTAG (the New and Emerging Respiratory Virus Threats Advisory Group), formally advises the Department of Health and Social Care, but has throughout collaborated closely with SAGE.

Formally, SAGE is a sub-committee of COBR where it is represented by Chris Whitty, as Chief Medical Officer, and Patrick Vallance, as Chief Scientific Advisor. COBR is the UK's primary decision-making body for dealing with emergencies and is supported by the Civil Contingencies Secretariat (a non-scientific body). It is at COBR that SAGE's scientific advice is pitted against other forms of advice from the economic, security, political, administrative, and diplomatic spheres. COBR (and SAGE) documentation explicitly attests to the need to draw on a wide spectrum of expertise regarding the impact of an emergency, not just of (natural) scientific or technical nature, but also economic and legal. Furthermore, to achieve the UK government's strategic objectives, it undertakes to not only seek scientific advice but to "apply risk assessment methodology and cost benefit analysis within an appropriate economic model to inform decision-making" [16, section 2.5].

The national framework for risk assessment identifies three aspects to the co-ordination of scientific advice. First is the assessment of the state of scientific understanding regarding the risk. The SAGE guidance document indicates that this is to be achieved through the analysis and modelling of existing data and assessment, the review and validation of existing research and, "where previous research is limited or non-existent", the commissioning of new research [6, p. 12]. Second is the delivery of recommendations on policy interventions, including "potential scientific and/or technical solutions that can remove or mitigate the risks and/or manage the impacts, and the pros and cons of these." [6, p. 11-12]. The final aspect is reflection on the state of scientific agreement and understanding, including reporting on "the degree of consensus ...; differences in opinion (i.e., are there differences in scientific / technical opinion and what are the sources of disagreements?); and the degree and cause of uncertainty" [6, p. 12].

The scope of advice SAGE is commissioned to provide is in some respects very broad, extending as it does to assessment of possible policy interventions and to the uncertainty around their impact. But in practice this advice has been restricted to impacts on public health (indeed just physical health) and then only on the scientific-technical aspects of them. Since SAGE does not consider the impact of policies on the other constituents of wellbeing, such as income, its recommendations must be regarded as carrying an implicit 'insofar as the sole concern is public health ...' qualification. Likewise, it does not seem to advise on crucial aspects of interventions such as whether people are likely to comply with imposed restrictions. SPI-B, which would be the natural source of advice on such behavioural factors, explicitly says that it is "not asked to comment, and has not commented, on what interventions are effective or when they should be triggered" [15, p.1].

COBR, in contrast, is required to evaluate policies in terms of their comprehensive outcomes. Hence, its decisions should be informed by an assessment of the impact of pandemic policies on *all* constituents of wellbeing, drawing on the full gamut of relevant expertise and not just that which is coordinated by SAGE. Since COBR's discussions are not a matter of public record in the way that SAGE's are, it is not possible to say where this expertise is sourced from (presumably the Treasury and other parts of the civil service), or how different advice is integrated. But there are two salient possibilities, neither of which are fully satisfactory.

The first of these is that COBR pits the policy assessments presented by different scientific advisors against each other - recommendations by SAGE health scientists based on a regard for public health against recommendations by economists, for instance – and favours one in the light of how COBR has prioritised its objectives at the time. This is the kind of picture that emerges from media accounts of the Treasury and Department of Health battling for policy supremacy and may explain the policy swings from lockdown to ‘Eat Out to Help Out’ as one group or another ‘wins’ the debate on priorities.

This approach is unsatisfactory for the simple reason that the effects of the pandemic cannot be identified and responded to in isolation. They interrelate, with policy choices for one constituent of wellbeing having consequential effects elsewhere. If people get ill, or frightened of falling ill, this will affect economic activity. Likewise, poor and financially precarious people are likely to risk infection (and infecting others) by going out to work, in order to retain income. So a policy that looks good from the perspective of, say, just public health may be sub-optimal when account is taken of its effect on comprehensive outcomes. It follows that one shouldn’t simply trade-off the policy recommendations made by different groups of experts, since all might be recommending policies that are sub-optimal from a comprehensive perspective. Trade-offs need to be made when evaluating comprehensive outcomes, not when comparing policies that seek to promote different constituents of them.

The second possibility is that COBR does its own integrated assessment of the impact of policies on comprehensive outcomes by developing ‘in-house’ models of the interaction between epidemiological, economic and behavioural factors as well as other variables. This too is unsatisfactory. Modelling the complex interactions between wellbeing factors and interventions on them requires expertise that COBR does not have. Recruiting such expertise would obviate the role of SAGE, thus undermining the independence of the scientific advice. There are good reasons to keep the scientific modelling of policy impacts as least partially separate from the process of weighing policy outcomes. The modelling must, of course, be directed at the issues of concern to policymakers; but once the scientific work begins, its results should be free from political influence or wishful thinking.ⁱⁱⁱ

Policies should be subject to integrated scientific assessment of their impact on comprehensive outcomes in support of, and prior to, policy making. Integrated assessment of this kind is routinely used in making climate policy for precisely the reasons that make it beneficial here: the interactions between different wellbeing constituents (e.g., health and economic outcomes) are complex and the policies need to be judged in terms of their impact on combinations of these constituents. Some attempts to conduct an integrated form of assessment of COVID-19 policies can be found in the academic literature (see, for instance, [1] and [2]). But since modelling of this kind is not routinely done at present within the government, SAGE should commission it. This will require a wider range of expertise to be involved in SAGE and will shift the focus away from purely health outcomes. And it is SAGE, not COBR, that should be authoritative on the question of the impact of policy choice on comprehensive outcomes.

3. Uncertainties

In reaching a judgement as to what policies to adopt in response to the COVID-19 pandemic, decision makers must deal with uncertainty that is both severe and multifaceted. Much of this is empirical, concerning the impact of policies on the comprehensive wellbeing outcomes that they target. But policy makers also face evaluative uncertainty, and hence legitimate disagreement, in the assessment of the desirability of these outcomes. A crucial question is how, on the one hand, both these uncertainties should be represented, measured, and communicated and how, on the other, policy making should be sensitive to them.

Although the focus of this section is on empirical uncertainty, let's start by looking briefly at evaluative uncertainty. Comparisons of the comprehensive outcomes of potential pandemic policies involve difficult ethical judgements about how to trade-off outcomes of different kinds (such as health and income) for individuals with different characteristics (such as age). Lockdowns benefit the elderly more than the young, for instance, and reduce people's risk of death at the expense of their expected income. These trade-offs are expressed in the kind of cost-benefit analyses that COBR is supposed to use by finding monetary equivalents for gains and losses to individuals in all dimensions of wellbeing. However, how the monetary value of such gains and losses should be determined is controversial. How does one attribute monetary value to a percentage reduction to the risk of death for an individual? Whatever figure is adopted will have significant implications for the assessment and selection of policies [1]. Despite this, there has been no transparency in the UK about which figure is being used or, indeed, more generally about how these trade-offs are made. In cases where SAGE has made direct policy recommendations, e.g., for school closure [16], there is no evidence that COBR instructed them on how to make these trade-offs. They were either not considered (a failure of responsibility) or made by the scientists (which is democratically inappropriate, as we said above). As a matter of priority this situation should be rectified, since **comprehensive policy assessment is impossible without an informed evaluation of potential policy outcomes, made by democratically empowered authorities, that recognises any uncertainty/disagreement and its implication for policy choice.**

Let us turn now to empirical uncertainty. To provide advice on the impact of the pandemic and of potential policy responses on comprehensive wellbeing outcomes, SAGE draws on the projections of causal models developed both 'in-house', by SPI-M-O and SPI-B, and more broadly by the scientific community. Both the modelling of the core epidemiological system and that of the effects of policy interventions on it are subject to considerable uncertainty. Some of this uncertainty is represented 'internally' by projecting the values of variables describing the system and communicating these as either intervals of values (e.g., that the reproduction rate R is between 1.3 and 1.6) or features of probability distributions over them (e.g., that the 95% confidence interval for R is 1.2—1.8). But some of it is 'external', in that it concerns the adequacy of the models themselves.

Consider, for instance, the SIR models that are the workhorses of epidemiology. In these models the population is divided into four buckets: susceptible to infection, infected, dead, or recovered. At each point in time the number of individuals in each bucket is determined by the number of people previously in each bucket and parameters such as the reproduction rate (R) and the infection to fatality rate (IFR). In the early phases of the pandemic, estimates for all of these were unreliable. Lack of testing capacity, for instance, meant that estimates for the number of people infected had to be inferred from the numbers getting ill, without knowledge of the prevalence of asymptomatic infection in the population. And while estimates have improved over time, there remains sufficient uncertainty about them that SPI-M-O's scenario modelling reports for SAGE still routinely begin with the disclaimer:

“The precise timings of peaks in infection and, in particular, demand on healthcare are subject to **significant uncertainty**. The scenarios are sensitive to initial conditions and any increase in the starting estimates of numbers of infections, hospitalisations, or deaths could lead to a larger peak.” [15, p. 1]

A second source of uncertainty concerns the adequacy of the causal analysis itself. Models cannot include every causal factor and there is always room for concern as to whether relevant ones have been excluded. Swamped health systems lead to deaths from causes other than COVID-19, an effect which was not accounted for by early models, including Imperial College's [7]. Nor was the endogenous effect of the spread of the virus on social distancing (e.g., scared people begin social distancing regardless of policy). In principle these effects can be modelled, but behavioural responses to the pandemic and the policy responses to it have multiple context-sensitive causes and are very difficult to model accurately. Furthermore, the fact that the proposed policies have rarely, if ever, been implemented in similar conditions means that policy models have very little data to calibrate against.

Standard ideal theory of science-based decision making prescribes a probabilistic representation of uncertainty combined with maximisation of expected benefit relative to these probabilities. SAGE and COBR guidance appear to be in tune with this in calling for cost-benefit analysis of policies and for advice on how likely different scenarios are. But SAGE's modellers have focused on generating scenarios based on different 'reasonable' assumptions about initial conditions and parameter values and have largely eschewed assigning probabilities to these assumptions. This is presumably a response to the fact that the prevailing level of scientific understanding is inadequate for them to make precise probability ascriptions with any confidence.

But how does one make policy choices based on a range of scenarios? At present, the government has not incorporated a range of scenarios into its decision-making process, instead selecting and reporting a particular scenario – the 'reasonable worst-case' – and basing decision making on the projections associated with it (and only it). In the next section, we evaluate this approach in more detail. But first we need to ask whether scientists should be selecting a single scenario to rely upon when a whole range of potential scenarios are consistent with the evidence. Selecting one makes life simpler for policy makers, but at the cost of denying them important information.

Firstly, consider how output values vary as a result of the assumptions built into each scenario. SPI-M-O captures one aspect of this by assessing how sensitive the projections associated with the reasonable worst-case scenario are to changes in varying the values of key variables. But scientists could also report the full range of values for an output variable generates by considering all plausible scenarios, e.g., by reporting that the value of R lies between 1.1 and 1.7, where these are the lowest and highest R-values present in any plausible scenario. Secondly, consider the level of scientific understanding underpinning the range of plausible scenarios. Scientists can make and express judgements about their confidence in different interval-valued projections depending on how broad the range of scenarios is with which they are consistent. For instance, they could report that their best estimate for R is 1.1 but that they have low confidence in this estimate, while they have medium or high confidence in R being in the interval between 1 and 1.2 (with these confidence judgements depending on how much evidence they have and how thoroughly the question has been investigated). Communication of such second-order confidence judgements is required by the IPCC's official uncertainty guidance [10] for instance (see also [9]).

Although some of this information is currently produced for SAGE and the resources exist for producing all of it, it does not seem to play any role in COBR's decision making. This is regrettable, since information about the characteristics of the set of plausible scenarios is useful to decision makers. Above all it tells them how *robust* the rationale for the decision is. A policy may be optimal relative to one scenario, for instance, but sub-optimal for many others. Or it may work well across a narrow interval of values for key variables, but not for a wider one. In such cases, decision makers have poorer grounds for confidence in their policy choice than in cases in which the choice is optimal across a larger range. So, when a wide range of scenarios cannot reasonably be excluded on scientific grounds, **scientists should communicate any characteristics and factors relevant to decision-making for the full range of possible scenarios**, and not just those associated with a single scenario.

4. Reasonable Worst-Case Scenarios

The concept of the reasonable worst-case scenario (RWCS) figures prominently in the output of scientific advisory committees and in the corresponding planning arrangements of government agencies. As government documentation emphasises repeatedly, the RWCS is "a generic representation of a challenging yet plausible manifestation of a risk. ... *It is not a prediction of what will happen*, rather an illustration of what we could reasonably expect to arise which is proportionate to use for preparation and planning purposes as a responsible government" [5, p. 4, emphasis added].

There are three questions to be addressed here. How is the reasonable worst-case to be defined? How can the worst outcomes be identified in uncertain conditions; in particular, regarding government policy? And why should this scenario, as opposed to ones worse or better than it, be the focus of scientific and decision-making attention?

The informal idea of a RWCS is relatively clear. We understand that the nature of the pandemic means that there are a range of possible outcomes within the different models that can be adopted. These outcomes can be (at least partially) ordered in terms of social/governmental preferences and values, and in terms of how plausible they are. Some scenarios are so implausible that they are irrelevant for decision-making purposes, with the RCWS being “the worst case once the high-impact low-likelihood manifestations of a risk have been discounted” [6, p.4]. SAGE documents don’t specify what counts as ‘low-likelihood’, but the National Risk Register declares that for a risk to be labelled as a RWCS it must have “at least a 1 in 20,000 chance of occurring ... and have an expected impact that reaches a minimum threshold (typically significant damage to human welfare in the UK)” [12, p. 69]).

Evidently, which scenario is identified as the reasonable worst-case will depend on what range of outcomes are considered and what the evaluative standards are used to order these outcomes. SAGE’s RWCSs are based upon a quite narrow set of characteristics, such as the number of infections, hospitalisations, ICU admissions and deaths over time. These are “worst cases” primarily in terms of public health outcomes, and specifically physical illness and mortality. They may not be worst case in terms of comprehensive wellbeing once the impact on people’s mental or financial wellbeing is included in the evaluation. But we argued before that when assessing the merits of a particular policy, regard should be had to all the significant consequences of adopting it, not just those relating to public health, a position apparently endorsed by the National Risk Register. So, the focus of scientific and decision-making attention on the public health worst-case scenario is not generally justified. **For RWCSs to be useful for policy evaluation they must be identified relative to comprehensive wellbeing outcomes.**

This leads to the second question, regarding how the reasonable worst-case scenario is to be identified. All the potential worst-case scenarios presumably meet a minimum threshold for damage, but it is not clear how the plausibility or likelihood of them is determined by SAGE. There are two difficulties here. The first, as we have seen, is that each scenario involves a host of assumptions about the magnitudes of key epidemiological and behavioural variables whose probability is extremely difficult to estimate. The second difficulty is more fundamental. What future public health outcomes are plausible, and how probable they are, depends on what policies the government adopts. How then are scientific advisors to make projections if government policy has yet to be decided? They cannot sensibly treat government policy as a random variable for which probability distributions can be estimated since the projections emanating from the modelling are supposed to be a factor in determining what that policy should be. **If policy follows the science, then science cannot pre-empt policy.**

To escape this impasse, SPI-M-O explicitly involves the Cabinet Office in deciding these model parameters, something that depends on what behavioural and social interventions (BSIs) will be made. In particular, the “values for R chosen after the easement BSIs have been agreed, ... in collaboration with SAGE and the Cabinet Office Civil Contingencies Secretariat” [15, p.6]. It follows that the reported scenario is *not the reasonable worst-case at all*. It is the worst-case, on the assumption that key variables (including R) correspond to the agreed values. The ‘true’ reasonable

worst-case, on the other hand, depends on how plausible/likely it is that these variables take one or another value, something that is sensitive not only on which BSI is chosen but how successfully it is implemented. (This difference, between the agreed RWCS and the true RWCS, perhaps explains why actual outcomes have frequently been worse than those projected by the former, a fact that would seem to cast doubt on the whole exercise.)

This leads to the third question: What is the rationale for focusing on RWCSs when a broad range of scenarios are considered possible?^{iv} It is helpful to distinguish the ‘upstream’ use of RWCSs to settle basic government policy from their ‘downstream’ use in the planning of organisations pursuing their more local goals in light of government policy. In the latter context, organisations may want to know what possibilities they should plan for. Consider for example a regional NHS manager who needs to prepare for a surge in the number of COVID-19 patients. To avoid services being overwhelmed she will want to know how much capacity to create by, for instance, cancelling routine operations and sectioning off facilities for COVID-19 patients. Although there are obvious costs to preparing for the worst-case, outcomes associated with the overwhelming of services may be sufficiently bad as to justify incurring them.

No such rationale can be offered for the use of RWCSs in upstream policy decision making. Since policy choices influence key epidemiological variables and the R-number in particular, any plausible criterion for identifying the RWCS will be sensitive to the policy adopted by the government. It follows that a policy choice cannot be rationalised by it being the best response to the RWCS. **If the RWCS is an outcome of policy choice, it cannot be assumed to be the case for the purposes of choosing a policy.** This would be to confuse the state of the world within which a policy is adopted from the consequence of adopting it.

5. Conclusion

The UK’s risk assessment framework mandates policy making based on the advice of scientific experts. That it does so, and in a manner that is fairly transparent, is something to celebrate. But aspects of the framework, and how it has been implemented in the pandemic, are unsatisfactory and steps should be taken to rectify them.

In the first place, the UK’s pandemic policies are not being chosen on the basis of a comprehensive assessment of their impact on individual wellbeing. We identified two aspects to this: that the integration of advice regarding different kinds of impacts is not being conducted by SAGE, but rather—if at all—by COBR, and that the evaluation of comprehensive outcomes, and the trade-offs that are involved, are not being made transparently and sometimes not by the right parties.

In the second place, the management of the uncertainty around the pandemic and policy responses to it is focused on the identification and modelling of reasonable worst-case scenarios. But a RWCS cannot coherently be used as a justification for a ‘upstream’ policy decision as it is also

a consequence of that policy. Furthermore, the focus on a single scenario is unjustified for two reasons.

First, even if we accept that in situations of severe uncertainty it is not possible to attach probabilities to different possible values for model parameters and initial conditions, there is decision-relevant information contained in the full set of reasonable scenarios that these models generate. There is, for instance, a policy relevant difference between a situation in which all reasonable scenarios project that R is in the interval $[1,3]$ and one in which they all project a magnitude in $[2,3]$, even if the RCWS is the same in both.

Secondly, scientists ought to say something about their confidence in the judgement that values for variables like R lies within a reported interval. Policy makers can use this second-order information to determine how cautious they need to be. The growing literature on decision making under severe uncertainty (or ‘ambiguity’ as it is usually termed in this field) is a rich resource to be mined in this regard (see [8] for a survey, [4] for an application of a confidence-sensitive decision rule to climate policy and [13] to health care).

These conclusions are relevant beyond the UK, in countries with similar frameworks for eliciting and using scientific evidence or which make use of worst-case scenario planning. For example, Sweden’s Public Health Authority uses worst-case scenarios in its modelling and planning, even though the relationship between experts and decision-makers is quite different to the UK. But detailed comparative assessment of national infrastructures for the use of science in emergency planning, and the extent to which they meet the desiderata we have argued for, is beyond the scope of this paper.

Acknowledgements

We are grateful to Krister Bykvist, Liam Bright, Orri Stefánsson, Alex Voorhoeve, and Kai Spiekermann for comments on an earlier draft.

References

1. Adler, M., R. Bradley, M. Ferranna, M. Fleurbaey, J. Hammitt and A. Voorhoeve “Assessing the Wellbeing Impact of the COVID-19 Pandemic and Three Policy Types: Suppression, Control, and Uncontrolled Spread”, T20 Policy brief, 2020, https://www.g20-insights.org/policy_briefs/assessing-the-wellbeing-impacts-of-the-covid-19-pandemic-and-three-policy-types-suppression-control-and-uncontrolled-spread/
2. Alvarez, F., D. Argente and F. Lippi “A Simple Planning Problem for COVID-19 Lockdown”, March 2020, NBER WP 26981
3. Birch, J. “Science and Policy in Extremis: the UK’s initial response to COVID-19”, 2020, retrieved from: <https://philpapers.org/rec/BIRSAP-4>

4. Bradley, R., C. Helgeson and B. Hill "Climate Change Assessments: Confidence, Probability and Decision", *Philosophy of Science* 84(3): 500-522, 2017.
5. Cabinet Office "Reasonable Worst-Case Scenario for borders at the end of the transition period on 31 December 2020", September 2020. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/920675/RWCS_for_our_borders_FINAL.pdf
6. Civil Contingencies Secretariat "Enhanced SAGE Guidance: A strategic framework for the Scientific Advisory Group for Emergencies (SAGE)", October 2012. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/80087/sage-guidance.pdf
7. Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., van Elsland, S., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P. G. T., Walters, C., Winskill, P., Whittaker, C., Donnelly, C. A., Riley, S., & Ghani, A. C. "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand", 2020. Retrieved from: <https://www.gov.uk/government/publications/impact-of-non-pharmaceutical-interventions-npis-to-reduce-covid-19-mortality-and-healthcare-demand-16-march-2020>
8. Gilboa, I. and M. Marinacci "Ambiguity and the Bayesian paradigm". In *Advances in economics and econometrics, tenth world congress, Volume 1*, 2011.
9. Helgeson, C., Bradley and B. Hill "Climate Change Assessments: Confidence, Probability and Decision". *Philosophy of Science*, 84 (3): 500-522, 2017
10. Mastrandrea, M. D., C. B. Field, T. F. Stocker, O. Edenhofer, K. L. Ebi, D. J. Frame, H. Held, E. Kriegler, K. J. Mach, P. R. Matschoss, G.-K. Plattner, G. W. Yohe, and F. W. Zwiers "Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties." Technical report, Intergovernmental Panel on Climate Change (IPCC), 2010.
11. National Risk Register 2017. Retrieved from: <https://www.gov.uk/government/publications/national-risk-register-of-civil-emergencies-2017-edition>
12. Roussos, J., R. Bradley, and R. Frigg "Making Confident Decisions with Model Ensembles". *Philosophy of Science*, forthcoming 2020.
13. Rowe, T and A. Voorhoeve "Egalitarianism under severe uncertainty", *Philosophy and Public Affairs* 46 (3):239-268, 2018
14. Scientific Pandemic Influenza Group on Modelling [SPI-M-O]. (2020). Covid-19 reasonable worst-case scenario planning, 21 May 2020. Retrieved from: <https://www.gov.uk/government/publications/spi-m-o-covid-19-reasonable-worst-case-planning-scenario-21-may-2020>
15. Scientific Pandemic Influenza Group on Behaviour [SPI-B]. (2020). The role of behavioural science in the coronavirus outbreak, 14th March 2020. Retrieved from: <https://www.gov.uk/government/publications/spi-b-the-role-of-behavioural-science-in-the-coronavirus-outbreak-14-march-2020>

16. UK Government Concept of Operations, 2013, HM Government:
<https://www.gov.uk/government/publications/the-central-government-s-concept-of-operations> Scientific Advisory Group for Emergencies [SAGE]. (2020). Seventeenth SAGE meeting on Covid-19, 18 March 2020. Retrieved from:
<https://www.gov.uk/government/publications/sage-minutes-coronavirus-covid-19-response-18-march-2020>

ⁱ <https://www.gov.uk/government/organisations/scientific-advisory-group-for-emergencies/about>

ⁱⁱ <https://www.gov.uk/government/publications/scientific-advisory-group-for-emergencies-sage>

ⁱⁱⁱ As Birch discusses in his analysis of the SAGE minutes, the relationship between SAGE and COBR changed during the summer of 2020 in a way that appears to make it more susceptible to wishful thinking and political influence, as well as worsening some of the issues discussed below in the section on reasonable worst-case scenarios. Birch quotes SAGE minutes from September 2020 stating that the RWCS was *agreed upon* by COBR and builds in which policy measures will take place [3].

^{iv} See [13] for a criticism of the use of worst-case scenarios in health decisions.