

Three ways in which pandemic models may perform a pandemic

Philippe van Basshuysen, Lucie White, Donal Khosrowi and Mathias Frisch

Abstract. Models may not only represent, but also *influence* their targets in important ways. While models' ability to influence outcomes has been studied in relation to economic models, often under the label "performativity", we argue that this phenomenon also pertains to epidemiological models, such as those used for forecasting the trajectory of the COVID-19 pandemic. After identifying three ways in which a model by the COVID-19 Response Team at Imperial College London (Ferguson et al. 2020) may have influenced scientific advice, policy, and individual responses, we consider the implications of epidemiological models' performative capacities. We argue, first, that performativity may impair models' ability to successfully predict the course of an epidemic; and second, that it provides an additional respect in which these models can be successful, namely by changing the course of an epidemic.

Keywords. COVID-19; epidemiological models; performativity; prediction; success; model evaluation.

1. Introduction

Philosophers are increasingly scrutinising epidemiological models that have been used to forecast the trajectory of the COVID-19 pandemic (Horner and Symons 2020; Nguyen and Frigg 2021; Winsberg et al. 2020; Schroeder 2021; van Basshuysen and White 2021). The success of these models is often evaluated in terms of their predictive accuracy (Friedman et al. 2020). We argue that there is an additional dimension along which epidemiological models can be successful, namely in *influencing* the course of an epidemic. Our argument draws on the concept of "performativity", which economic sociologists and philosophers of economics have used to describe the phenomenon that economic models may not only represent parts of the economy, but can also influence the behaviours of the economic systems modelled (Callon 1998; MacKenzie 2006; Guala 2007; Mäki 2013; Boldyrev and Ushakov 2016). We understand the notion of performativity in a broad way; models *perform* by changing important properties of the phenomena they are representing (cf. Mäki 2013).¹ This notion has a long history in philosophy

¹ Mäki himself is critical of attempts to characterize this type of phenomenon in the context of economic modelling in terms of Austin's notion of performativity, arguing that Austin's notion is concerned with acts being *constituted* by certain utterances. By contrast, according to Mäki, economic theorizing may have *causal* consequences for economic activity without being constitutive of it. While we do not want to engage in a dispute on whether Austin's notion is apt in this context, we wish here to draw attention to the fact that, like economic models, epidemiological models can have causal effects on the systems modelled.

of science (see Buck 1963), and economics and the social sciences (see Henshel 1993), often referred to as “reflexive prediction” (Buck 1963; Henshel 1993). Predictions (or “prophecies”) can be “self-fulfilling” or “self-defeating” or “-frustrating”, understood broadly to mean that the dissemination of the prediction is causally relevant in either bringing about the event predicted, or its non-occurrence (Buck 1963). Whereas earlier accounts tended to focus on the negative aspects of this phenomenon (particularly the difficulties it poses for accurate prediction), contemporary work on performativity also emphasises potential positive aspects, particularly, how understanding performativity allows us to intervene constructively in the economy (Guala 2007).

Despite this sustained discussion, the question of whether epidemiological models might be performative has thus far been neglected in philosophical studies of these models.² We aim to make a first step towards closing this gap by arguing, first, that performativity is a phenomenon pertaining to epidemiological models, and second, that this has important implications for how these models should be regarded and evaluated. We proceed as follows: after sketching some important features of epidemiological models, we focus on a model that was used by researchers at Imperial College London (ICL) to forecast the pandemic in the UK and the US (Ferguson et al. 2020). Using this example, and focused particularly on its use in the UK, we identify three ways in which models may be performative: (i) models may influence policy advising; (ii) models may affect policy-makers’ responses to an epidemic; and (iii) models may directly influence individual responses to an epidemic. While (i) and (ii) are well-documented in the case of the ICL model, (iii) is more speculative.³ Having identified the ways in which epidemiological models can be performative, we elaborate how their performative aspects may exacerbate the task of successfully forecasting the course of an epidemic. At the same time, however, we contend, their performative abilities provide an additional way in which these models can be successful, namely by changing the course of an epidemic.

2. Epidemiological models

Epidemiological models are among the instruments policy makers have used (and continue to use) in making policy decisions during the COVID-19 pandemic. These models provide highly idealized representations of how the virus will spread through a population under a variety of assumptions.

² Some social scientists have described epidemiological models as performative (Rhodes et al 2020), but their focus differs from ours, as their main concern is that we should adapt these models to the changing circumstances.

³ The first two types of “performativity” we outline could apply to any model that is used to inform policy, whereas the third sense is more distinctive and methodologically interesting, and it is on this notion that most theorists focus (see e.g. the articles in MacKenzie, Muniesa and Siu 2007). Thanks to an anonymous reviewer for pointing this out.

The simplest models are SIR or SEIR models, which assign members of a population to different subgroups, “the Susceptible”, “Infectious”, “Removed/Recovered”, or, in the case of SEIR models, also the “Exposed”. SIR and SEIR consist of several deterministic equations that describe how the proportional sizes of the different subgroups change with time, based on transition rates between these groups. This depends on a small number of parameters: the transmission rate, the recovery rate, and the latency period. SIR or SEIR models are macroscopic models that treat populations as a whole, and abstract from interactions among individuals and transmission events.

The model used by the Imperial College London to model the global evolution of the pandemic in different regions of the world is a SEIR model (Walker et al. 2020). By contrast, the “CovidSim” model discussed in the highly influential “Report 9” by the Imperial College COVID-19 Response Team, upon which we chiefly focus, is an individual-based model (Ferguson et al. 2020), in which individuals are assigned to different types of location where contacts occur—within the household, at school, in the workplace, and in the wider community. Transmission events through contacts among individuals at these locations and movements between locations are then modelled. CovidSim is a stochastic model containing a very large number of parameters, many of which are only poorly constrained by empirical data. A sensitivity analysis (Edeling et al. 2021) has shown that the model is particularly sensitive to variations in three parameters: The length of the latent period in which a patient has no symptoms and is not infectious; the delay to start case isolation; and, the ‘relative spatial contact rate given social distancing’ parameter, which captures the effectiveness of social distancing.⁴

3. Three performative aspects of the ICL model

(i) The ICL model performed scientific advising.

In a recent working paper, Jonathan Birch (2020) documents how the form in which scientific advice was provided in the UK evolved during the early stages of the pandemic, in the lead up to the introduction of the ICL model used in Report 9 (Ferguson et al. 2020a). He suggests that the Scientific Advisory Group for Emergencies (SAGE) – the group chiefly responsible for guiding the UK’s initial pandemic response – provided advice in three distinctive forms during March 2020. They began by providing advice in the form of what Birch refers to as “no unconditional recommendations”: that is, they presented a list of potentially effective means of pursuing a variety of goals that policy-makers might have, with no indication of which goals might be desirable or worthy of pursuit. As the situation worsened, by the 9th of March, SAGE shifted to

⁴ This last parameter in particular will be important in the following.

providing advice in the form of “disjunctive unconditional recommendations”; endorsing a specific end, but leaving a variety of means open for achieving this end. They did this by outlining a range of non-pharmaceutical interventions that they believed could, in some combination, achieve epidemic mitigation, and recommending that some subset of these measures be implemented (SAGE 2020a). On the 18th of March, shortly after SAGE had begun to discuss the ICL modelling paper, the form of advice shifted again, to a “single unconditional recommendation”: school closures are required as soon as possible to prevent the NHS from becoming overwhelmed (SAGE 2020b).

Birch traces this trajectory in order to argue that the latter form of advice, while perhaps not appropriate in normal situations, is appropriate in fast-moving crisis situations. But he also comments in passing on a possible reason for this final shift to advice in the form of a single unconditional recommendation. Namely, two days before issuing their single unconditional recommendation (on the 16th of March), the SAGE team discussed the ICL modelling paper for the first time. This paper displayed data about critical care capacity in the same graph as projections about demand for critical care in the absence of strict suppression measures (Figure 1), revealing, in stark form, the extent to which demand was predicted to outstrip supply, even with existing mitigation measures in place. Birch suggests that this might have been the first time that some SAGE members were made aware of the dramatic size of the mismatch. The ICL modellers also included a single unconditional recommendation in the report on the basis of their projections: “epidemic suppression is the only viable strategy at the current time” (Ferguson et al. 2020). It is plausible, then, that this model - particularly through the graphical representation of its results⁵ - was instrumental in shaping the form of the scientific advice provided by SAGE.

⁵ The ways in which graphical representations of models and their results may influence how these models are adopted is an important topic in the performativity literature; see, for instance, Pahl and Sparsam (2016) in the context of the IS-LM model.

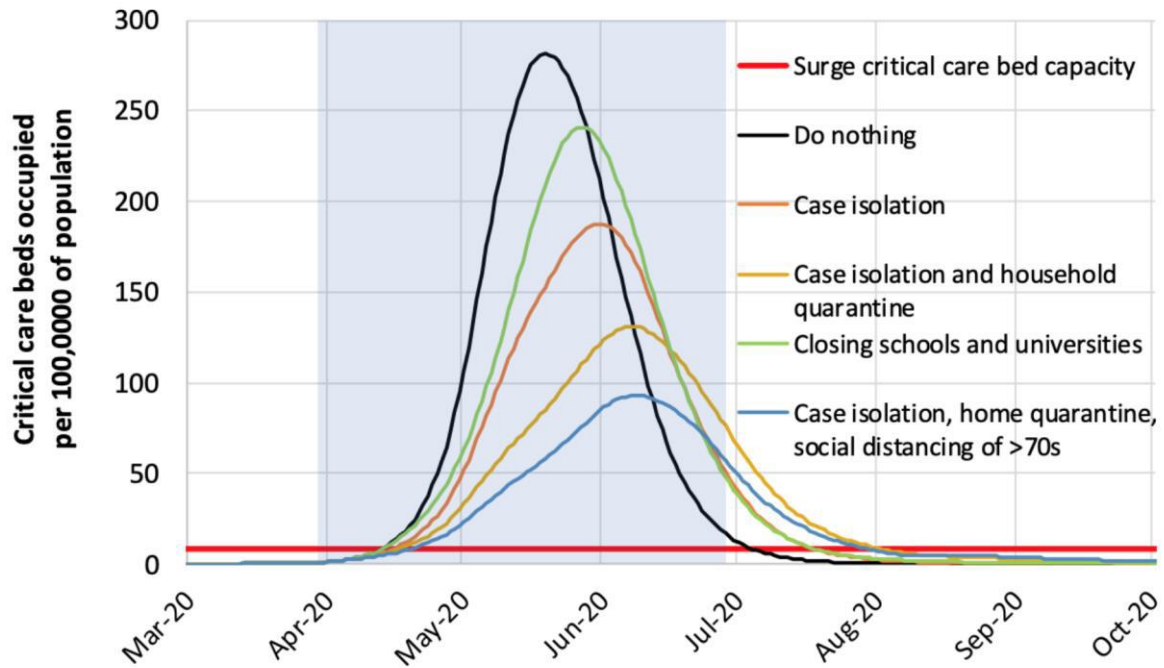


Figure 1: Projected demand for critical care beds under different possible mitigation-based policies is displayed together with surge critical care bed capacity; from Ferguson et al. 2020.

(ii) *The model performed policy-makers' responses to the pandemic.*

It is clear that the ICL model had a significant impact on public policy in the UK; immediately following the release of Report 9, on the 16th of March, the UK drastically altered (Bosely 2020) its previous, mitigation-based strategy (DHSC 2020) and almost all of the ICL team's recommended measures were implemented in the following days (see van Basshuysen and White forthcoming). This may have further led, however, to some noteworthy interactions between models projecting the trajectory of the pandemic and providing policy advice on their basis, and to the response of policy makers. The ICL model and models like it, which project the impact of various public policy measures, are designed to provide policy guidance. Their usefulness depends on their predictive performance. But the implementation of the recommendations of these modellers steers the course of the epidemic away from their reference forecasts. This can lead to claims that the model is "too pessimistic", seemingly based on the idea that the actual numbers of those who are infected, hospitalised, or die, do not resemble the model's projected numbers, where these numbers project what will happen in the absence of any interventions or behavioural changes, but where mitigation measures have been implemented (see e.g. St. Onge and Campan 2020). We might regard this as a distinctive instance of *counterperformativity* (see Mackenzie 2006) – the implementation of the recommended policies serve to steer the real world

outcome away from the reference forecast, but notably, in the case of the ICL model, this was the exact effect that the modellers were hoping to achieve.⁶

Joseph Friedman et al. draw attention to another problem in their evaluation of the predictive performance of various international COVID-19 mortality forecasting models. They suggest that transmission-based epidemiological models tend to overestimate transmission in their reference forecasts, because prolonged exponential growth in transmission rarely continues unabated – governments (and/or individuals, see below) will change their behaviour in response to deteriorating circumstances--and, we contend, in response to updated modelling projections. The accuracy of modelling predictions and projections could be improved, they argue, by taking such endogenous factors into account (which, as they note, are often explicitly included in economic analyses). In addition, they suggest that models (such as the ICL models) that aim to predict the impact of various public policy measures would be better evaluated by looking at the error in the difference between the reference forecast and the forecast for the implemented policy. Friedman et al. maintain, however, that these counterfactuals can be difficult to gauge. Even when comparing the projection concerning the implemented policy to its actual effects, the effects of the policy must be somehow distinguished from other factors (2020). It should however be noted that economists have made progress in recognising and incorporating endogeneity into COVID-19 modelling, which, among other things, can assist in clarifying the role of government policy in disease dynamics (see Avery et al. 2020).⁷

The belief concerning the inaccuracy and overly pessimistic nature of these models has potential further follow-on effects for the implementation of the policy advice on the basis of further models. It is possible that the scepticism and controversy surrounding the ICL projections was one of the factors that contributed to a reluctance to implement a further, short term, “circuit-breaker” lockdown when a second wave was unfolding in the UK in autumn 2020 (see Sayers 2020), which SAGE called for with increasing urgency (Stewart and Sample 2020), on the basis

⁶ Although this phenomenon has led to criticism of such models for generating inaccurate predictions, and comparison of the reference forecast to the actual outcome is used to evaluate the predictive accuracy of these models (see e.g. Friedman et al. 2020), it should be emphasised that the projection here is contingent on “no measures being implemented” (i.e. this is a conditional prediction). It is helpful here to distinguish predictions, which generate statements about future states of the actual world, from projections, which generate statements about counterfactual outcomes, conditional on some set of counterfactual assumptions. The performative effect of the model consists in influencing which member of a set of projections ends up most closely representing the trajectory of the epidemic in the actual world and does not compromise its predictive accuracy. While the model steers the trajectory of the epidemic further away from the trajectory in the reference projection, this does not undermine the predictive accuracy of the model. Thus, this does not constitute the type of “self-defeating prediction” that is the focus of much attention in the performativity literature, where, for example, the prediction that an imminent disaster *will happen* prevents this disaster from happening and thus undermines the accuracy of the prediction (see e.g. Henshel 1993). Thanks to an anonymous reviewer for pressing this distinction.

⁷ Thanks to an anonymous reviewer for making us aware of this work.

of mathematical modelling projections (SAGE 2020c). The refusal to implement this advice constituted a significant break between scientific advisors and policy makers in the UK (Stewart and Sample 2020).

(iii) *The models may have performed individual responses to the pandemic.*

While epidemiological models have influenced individual behaviour through scientific advising and policy making, these models may also have been performative in a more unmitigated way; changing individual behaviour directly as a result of the information conveyed by the model. In the past year governments have introduced a variety of measures to reduce and restrict contacts among people with the aim of reducing the spread of COVID-19. The effectiveness of many of these measures, however, depends to a significant extent on how the public interprets the threat presented by the pandemic, and their willingness to comply with any measures implemented. And this in turn depends, at least partly, on how the widely discussed epidemiological predictions concerning infection rates or the reproduction rate R_0 , derived from epidemiological models, are understood and interpreted by the public.

The success of any policy measure is predicated, at least in part, on the degree of public compliance with the measure in question. Whether a policy restricting all social gatherings to the members of one household plus one additional person—a restriction in place in Germany as of the date of writing—can contribute to a reduction in R_0 depends on the extent to which people abide by this restriction. A particular model may not distinguish between (i) the institution of a policy and (ii) degrees of compliance with this policy, simply assuming that instituting a given policy will result in universal compliance. But to what extent this idealization is adequate will depend on the degree of compliance, which in turn will at least partly depend on how the models' predictions are understood and taken up by the public. The success of many measures, furthermore, goes beyond the issue of compliance. Measures such as the abovementioned restriction on social interactions are compatible with a wide range of social networks, including quite extensive networks. The extent to which such measures slow the spread of infection is thus partially dependent on how individuals interpret a policy and how they implement it in adjusting their interactions. There are many different ways, for example, of following the 'one-additional-person'-rule: I might meet with a large number of different friends individually night after night; or I might radically cut down on my social contacts. Both behaviours comply with the rule, but the resulting social networks end up looking quite different. Thus not only degrees of compliance, but also how a given policy is interpreted and implemented through individual

behaviour, is contingent on how serious individuals take the threat presented by the pandemic to be.

This is also the case when explicit policy guidance is lacking – for example, the number of infections that result from interactions in the workplace (which is one of the types of interaction explicitly modelled in the ICL model) will depend to some extent on how individuals adjust their behaviour in response to the model’s predictions. The predictions of epidemiological models and how these are understood by the public can thus play an important causal role in changing properties of the pandemic—a connection that is further underscored by the empirically well-supported assumption that people in many places reduced their contacts in advance of the implementation of social distancing policies (e.g. Friedson et al. 2020; Sears et al. 2020). In part, this may have been due to new information regarding the pandemic, including model projections, which have figured prominently in news stories about the pandemic (e.g. Boseley 2020).

In short, our modelling efforts interact with the very quantities and properties we are trying to model and can thereby change them: epidemiological models and the quantities modelled by them exhibit what Ian Hacking has called a “looping effect” (1999, 105). A particularly vivid example of this in the case of the CovidSim model is given by the parameter representing the effectiveness of social distancing measures. As we have mentioned above, the model’s predictions are especially sensitive to changes in the value of this parameter. Yet the effectiveness of social distancing measures depends in turn on how the model’s results are communicated to and understood by the public. For instance, when people are provided with this information through some kind of graphical representation, such as the graph in Figure 1, depicting how the demand for critical care beds could greatly exceed capacity, this might make the urgency of the situation particularly salient, and it might lead to different reactions than if this information had been conveyed through other means.

4. The role of performativity in model evaluation and appraisal

Here, we wish to highlight two important ways in which thinking about epidemiological models’ performativity can improve our understanding of how to evaluate these models. We consider, first, how performative aspects may compound the difficult task of successfully forecasting the course of an epidemic and, second, how they might be assessed in light of their ability to change the course of an epidemic.

Epidemiological models serve dual purposes: apart from their epistemic purpose of forecasting the course of an epidemic, they also serve the practical purpose of informing and guiding policy making. These epistemic and practical purposes go hand in hand: on the basis of forecasts, policy-makers can choose policies that are likely to prevent unwanted outcomes. However, as we argued above with respect to some of the ICL's models, this may steer away the course of the epidemic from reference forecasts, which do not incorporate policy or other changes. If a modelling study does not include a range of scenarios reflecting different combinations of policies (some of which resemble those that will actually be implemented), some of the estimated parameters may turn out to be inaccurate due to behavioural changes (e.g. levels of mobility), and the forecasts derived from the model may thus be inaccurate, too. There is thus a way in which a model's performative capacities to shape policy making can impair the accuracy of its forecasts.

To be sure, many of the models that have been used to forecast the course of the COVID-19 pandemic, including the ones we have focused on here (Ferguson 2020; Walker 2020), do include a range of scenarios in which different combinations of policies are implemented. This allows projections that better resemble the actual policy trajectories followed in a country, and models' predictive accuracy can accordingly be assessed by comparing these projections with observed outcomes (van Basshuysen and White 2021; forthcoming). However, gauging the accuracy of even these more pertinent scenario forecasts can face significant limitations due to the fact that most models abstract from endogenous behavioural responses, such as when individuals engage in social distancing even before government-issued directives take effect, adopt more rigorous hygiene and personal protection protocols, cancel travel plans, etc. Such behavioural responses are poorly understood, difficult to anticipate (Friedman et al. 2020), and thus difficult to incorporate into models (which is perhaps why extant models make no attempts to include them). For instance, Report 9, described in the previous section, forecasts epidemic trajectories conditional on combinations of policies, but not on the spontaneous behavioural responses of the population, such as increased social distancing prior to a policy coming into effect, or in addition to it. As we noted above, such endogenous behavioural responses may be partly driven by the performative abilities of models. But if endogenous responses influence the course of the pandemic, this may distort the forecasts by impairing the accuracy of important parameter values.⁸ Furthermore, even sincere efforts at estimating these parameters may be

⁸ It should be noted that not all distortions due to spontaneous behavioural changes are the effect of performativity being in play. A distinction might be drawn here between individuals adapting their behaviour in response to model predictions, or in response to other kinds of information, such as current prevalence, which may accrue independently of the models in question. It might be difficult, however, to evaluate which mechanisms are in play in

limited, as behavioural responses can change over time (such as individuals becoming tired of social distancing directives and departing from recommended behaviours) – so there remain epistemic challenges even if one endeavours to conscientiously estimate crucial empirical parameters, and re-parameterize models on an iterative basis (as is done for some of the models we discuss). There are thus limitations to models’ predictive abilities raised by models’ performativity.

Our second point follows on the heels of the first: while the performative aspects of epidemiological models can hamper their ability to successfully forecast the course of an epidemic, they may nevertheless inform our judgments concerning whether these models have been useful in a second way, namely, by changing the course of an epidemic. Here, we emphasise that models’ performativity may figure as an important criterion for appraisal beyond predictive accuracy.

When models are performative, their dual roles as epistemic instruments (that predict the evolution of a pandemic) and as practical instruments (that inform policy-making through reference and scenario forecasts) can make it difficult to evaluate them with respect to each purpose individually. The performative effects that these models have on policy and individual behaviour can drive outcomes away from the models’ predictions. When the predictions of models are undermined, moreover, this seems to undermine their epistemic credentials, making them unsuitable as a basis for public policy or individual behavioural response. When we focus only on the predictive capacities of such models, we seem to be in a bind here - the more that they steer policy and behaviour, the less suitable they seem as a basis for steering policy and behaviour (unless they endogenize behavioural response - although, as we have just pointed out, this can be difficult).

One way out of this bind is to move model appraisal beyond focusing on models’ predictive performance alone. We thus suggest that both predictive and performative capabilities should be considered side-by-side when appraising epidemiological models, which can proceed by evaluating some forms of performativity as a desirable attribute (and others as undesirable) in all-things-considered assessments of what models have done for us, epistemically and practically. Following a Kitcher-style view that understands and evaluates science through the roles it plays in, and partly *for*, society, and where those roles often extend beyond purely epistemic ones, we

a given case as it may be unclear whether individuals’ sources of information include model results. For our point, that performativity may exacerbate the task of forecasting, it suffices that individual behaviour sometimes responds to model results, which should be uncontroversial, given that these results and their graphical representations have figured prominently in public perception, as we argued above. We thank an anonymous reviewer for pointing this out.

emphasise that a more comprehensive assessment of what epidemiological models have done for us (and whether they did so well) should consider their performative aspects, too.

To use an analogy, consider a medical doctor telling her patient that they are likely to die of a heart attack within the next decade if they continue smoking, drinking, and doing little exercise. Such a prediction can be performative in helping the patient change their habits, thus escaping the unwanted prediction. Likewise, epidemiological models can have the performative ability to change important aspects of the phenomena they seek to capture. While this can adversely affect their predictive abilities, this does not necessarily imply that a model's suitability, adequacy, or usefulness is diminished, if we consider its performative impact to be a potential virtue. Just as we should not think that a doctor is unqualified because she cannot tell whether a patient would indeed have died of a heart attack counterfactually, or because she couldn't accurately forecast how much longer, exactly, a patient would live under a changed exercise regime, we suggest that it is unhelpful to assess the utility of epidemiological models based on their predictive abilities alone. Epidemiological models have been a crucial resource for informing and justifying policy interventions, and may have contributed to shaping both the public's understanding of the pandemic, and their behavioural response to it. The ability to make such performative contributions can be understood as a desirable feature of these models, even if models' performativity can diminish their predictive prowess.

It is important to stress, however, that not all forms of model performativity are desirable. We might think, for instance, that some performative successes of epidemiological models can also have unwanted side-effects, such as when some members of the public begin to take the difference between observed outcomes (such as cumulative mortality) and widely publicized pessimistic reference scenarios as reason to call the severity or reality of the phenomena being modelled into question. One might also worry that by choosing particular outcome variables as key modelling targets, such as critical care demand or cumulative mortality, models have been performative in making these outcomes relatively more salient than various economic, societal, or psychological endpoints, and that models might thereby have contributed to steering policy and public response in ways that led to suboptimal societal outcomes.⁹ We do not take a stance on which performative aspects are desirable or undesirable here, but rather highlight that it is important to take both kinds into account.

So how, exactly, are we to take models' performative aspects into account? Unfortunately, numerous complications arise when detailing what role performativity might play in model

⁹ We thank an anonymous reviewer for raising this concern.

appraisal. Should it play an evaluative role, providing a criterion by which the merits or deficiencies of models can be judged? Or should it be prescriptive; recommending the construction, selection, or use of models with certain performative abilities? The latter view seems highly problematic: we do not think that models can (or should) be made better by being made more performative, even if particular forms of performativity are uncontroversially socially desirable. Deceitfully meddling with a model in order to steer policy or behavioural response would raise significant concerns about illegitimate value-influences, and could severely undermine the epistemic credentials of models. Models are, and should remain, epistemic instruments in the first instance, and while some value-influences are rarely avoidable, model construction and use should proceed in epistemically responsible ways (see Elliott and McKaughan 2014; Parker and Winsberg 2018 for discussions of the roles of non-epistemic values in model construction and use).

To allay such concerns, we need principles that allow us to acknowledge models' performative aspects, good and bad, while maintaining a commitment to the idea that models' epistemic functioning should remain untouched by considerations about their performativity. Unfortunately, formulating such principles, and providing a detailed account of how exactly we should weigh or integrate models' epistemic and performative aspects, involves various important and non-trivial questions that cannot be addressed here. These include: on which outcomes should we focus when gauging models' performative effects? How should we weigh considerations about models' epistemic and performative aspects? Should epistemic considerations enjoy lexical priority, or are there cases where performative virtues may compensate for epistemic shortcomings? Should all performative aspects be considered in model appraisal, or are there cases where we should distinguish between the model-as-such, the practice of model use, and issues of science communication about model outputs, where evaluations of performativity should pertain to the latter two but not the former? We suggest here, for now, only the following provisional safeguard: performativity should only play an evaluative, but not a prescriptive role in model appraisal. It may figure in judgments concerning whether the downstream performative aspects of models have been desirable or undesirable, but such judgments should not bear on decisions made at the stage of model construction, selection, or deployment – there should be no *wishful modelling*. Navigating the additional intricacies involved in formulating finer-grained principles is beyond the scope of this paper and must be addressed in future work.

Finally, we would like to highlight a rather different, and less problematic route to taking models' performativity into account at the stage of model construction: instead of aiming to build models with specific performative goals in mind, we can aim to build models that endogenize their own performativity. Social scientists, as early as the 1950s, have made efforts to develop analytic machinery that can help anticipate the performative effects of public predictions (Modigliani and Grunberg 1954; Simon 1954). It thus seems that a fruitful avenue for further research would be to investigate how the public's behavioural response to publicized model outputs as well as to policy measures could be captured by epidemiological models. Although we have noted the difficulties generated by attempts to incorporate endogeneity, there has been progress on this (see e.g. Avery et al 2020; Eksin et al. 2019). Doing so might not only be a promising way to help improve models' predictive capabilities, but would also answer criticisms that social scientists have been insufficiently involved in scientific advisory systems that inform policy. One important role to be played by social scientists (among many others) can be to directly contribute to the predictive endeavours undertaken by epidemiologists, by helping account for and anticipate the intricacies of the social and behavioural underpinnings that govern models' performativity.

5. Conclusions

We have argued that epidemiological models may have altered the course of the SARS-CoV-2-pandemic in three important respects: i) through shaping the form of scientific policy advice, ii) through shaping policy formation, and iii) through influencing individual behavioural responses to the pandemic. Existing philosophical literature on epidemiological models and extant model evaluation studies have, however, focused primarily on predictive accuracy, and these performative aspects, and their implications, have been largely neglected. Yet, we have proposed that it may be important to take the performativity of models into consideration. This is, first, because it can exacerbate the difficulties involved in accurately predicting the course of a pandemic, and second, because certain types of performativity might be regarded as a desirable feature of epidemiological models. It thus might be fruitful to consider, in addition to predictive accuracy, the ways in which we can assess the success of models based partially on their performative aspects, and the role these might play in our fight against the pandemic.

References

- Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fischer Ellison. 2020. "An Economist's Guide to Epidemiology Models of Infectious Disease". *Journal of Economic Perspectives* 34(4):79-104.
- van Basshuysen, Philippe, and Lucie White. 2021. "Bad data and flawed models? Fact-checking Winsberg et al.'s case against lockdowns". *LSE Philosophy Blog*, January 26. Accessed March 1, 2021.
- van Basshuysen, Philippe, and Lucie White. forthcoming. "Were lockdowns justified? A return to the facts and evidence". *Kennedy Institute of Ethics Journal*.
- Birch, Jonathan. 2020. "Science and policy in extremis: the UK's initial response to COVID-19". Working Paper. Accessed February 25, 2021. <https://philpapers.org/archive/BIRSAP-4.pdf>.
- Boldyrev, Ivan, and Alexey Ushakov. 2016. "Adjusting the model to adjust the world: constructive mechanisms in postwar general equilibrium theory." *Journal of Economic Methodology* 23(1):38-56.
- Boseley, Sarah, 2020. "New Data, New Policy: Why UK's Coronavirus Strategy Changed". *The Guardian*, March 16. Accessed March 1, 2021. <https://www.theguardian.com/world/2020/mar/16/new-data-new-policy-why-uks-coronavirus-strategy-has-changed>.
- Buck, Roger. 1963. "Reflexive predictions". *Philosophy of Science* 30(4): 359-369.
- Callon, Michel. 1998. *The laws of the markets*. Malden, MA: Blackwell.
- Department of Health and Social Care (DHSC). 2020. "Coronavirus action plan: a guide to what you can expect across the UK". March 3. Accessed March 1, 2021. <https://www.gov.uk/government/publications/coronavirus-action-plan/coronavirus-action-plan-a-guide-to-what-you-can-expect-across-the-uk>.
- Edeling, Wouter, Hamid Arabnejad, Robbie C. Sinclair, Diana Suleimenova, Krishnakumar Gopalakrishnan, Bartosz Bosak, Derek Groen, et al. 2021. The impact of uncertainty on predictions of the CovidSim epidemiological code. *Nature Computational Science* 1:128-135. <https://doi.org/10.1038/s43588-021-00028-9>
- Elliott, Kevin C. and Daniel J. McKaughan. 2014. "Nonepistemic values and the multiple goals of science". *Philosophy of Science* 81(1):1-21.
- Eksin, Ceyhan, Keith Paarporn, and Joshua S. Weitz. 2019. "Systematic biases in disease forecasting – The role of behavior change". *Epidemics* 27:96-105.
- Ferguson, Neil M., Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, et al. 2020. "Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand." *Imperial College COVID Response Team*. March 16. <https://doi.org/10.25561/77482>.
- Friedman, Joseph, Patrick Liu, Christopher E. Troeger, Austin Carter, Robert C. Reiner JR, Ryan M. Barber, James Collins, et al. 2020. "Predictive performance of international COVID-19

mortality forecasting models.” Version 5. *medRxiv Preprint*, November 19.

<https://doi.org/10.1101/2020.07.13.20151233>

Friedson, Andrew I., Drew McNichols, Joseph J. Sabia, and Dhaval Dave. 2020. “Did California’s Shelter-in-Place Order Work? Early Coronavirus-Related Public Health Effects.” NBER Working Paper 26992. <http://www.nber.org/papers/w26992>

Grunberg, Emile, and Franco Modigliani. 1954. “The Predictability of Social Events”. *Journal of Political Economy* 62(6): 465-478.

Guala, Francesco. 2007. “How to do things with experimental economics.” In *Do economists make markets? On the performativity of economics*, edited by D. MacKenzie, F. Muniesa and L. Siu, 128-162. Princeton, NJ: Princeton University Press.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Henshel, Richard L. 1993. “Do self-fulfilling prophecies improve or degrade predictive accuracy? How sociology and economics can both be right.” *The Journal of Socio-Economics* 22(2): 85-104.

Horner, Jack K., and John F. Symons. 2020. “Software engineering standards for epidemiological models.” *History and Philosophy of the Life Sciences* 42, 54. <https://doi.org/10.1007/s40656-020-00347-6>.

MacKenzie, Donald. 2006. *An engine, not a camera. How financial models shape markets*. Cambridge, MA: MIT Press.

MacKenzie, Donald, Fabian Muniesa, and Lucia Siu (eds.). 2007. *Do economists make markets? On the performativity of economics*. Princeton: Princeton University Press.

Mäki, Uskali. 2013. “Performativity: Saving Austin from Mackenzie.” In *EPSA11 Perspectives and Foundational Problems in Philosophy of Science*, edited by Vassilios Karakostas and Dennis Dieks, 443-453. Dordrecht: Springer.

Nguyen, James, and Roman Frigg. 2021. “What are ‘scientific models’, and how much confidence can we place in them?” *LSE Covid Blog*, February 10. Accessed February 27, 2021. <https://blogs.lse.ac.uk/covid19/2021/02/10/what-are-scientific-models-and-how-much-confidence-can-we-place-in-them/>.

Pahl, Hanno, and Jan Sparsam. 2016. “The IS–LMization of the General Theory and the Construction of Hydraulic Governability in Postwar Keynesian Macroeconomics”. In *Enacting Dismal Science*, I.A. Boldyrev, E. Svetlova (eds.), Palgrave Macmillan US, 2016. DOI 10.1057/978-1-137-48876-3_7

Parker, Wendy S., and Eric Winsberg. 2018. “Values and evidence: how models make a difference”. *European Journal for Philosophy of Science* 8:125-142.

Rhodes, Tim, Kari Lancaster, Shelley Lees, and Melissa Parker. 2020. “Modelling the pandemic: attuning models to their contexts”. *BMJ Global Health* 5:e002914. doi:10.1136/bmjgh-2020-002914.

- Sayers, Freddie, 2020. "The case for a circuit-breaker". *Unherd*, October 16. Accessed March 1, 2021. <https://unherd.com/2020/10/the-case-for-a-circuit-breaker/>
- Schroeder, S. Andrew. 2021. "How to Interpret Covid-19 Predictions: Reassessing the IHME's Model". *Philosophy of Medicine* 2(1). <https://doi.org/10.5195/philmmed.2021.43>
- Scientific Advisory Group for Emergencies (SAGE). (2020a). "Potential effect of nonpharmaceutical interventions (NPIs) on a COVID-19 epidemic in the UK". March 9. Accessed February 28, 2021. <https://www.gov.uk/government/publications/potential-impactof-behavioural-and-social-interventions-on-an-epidemic-of-covid-19-in-the-uk-9-march-2020>.
- Scientific Advisory Group for Emergencies (SAGE). (2020b). "Seventeenth SAGE meeting on Covid-19, March 18". Accessed February 28, 2021. <https://www.gov.uk/government/publications/sage-minutes-coronavirus-covid-19-response-18-march-2020>.
- Scientific Advisory Group for Emergencies (SAGE). (2020c). "Summary of the effectiveness and harms of different pharmaceutical interventions". September 21. Accessed March 1, 2021. <https://www.gov.uk/government/publications/summary-of-the-effectiveness-and-harms-of-different-non-pharmaceutical-interventions-16-september-2020>
- Sears, James, Villas-Boas, J. Miguel, Villas-Boas, Vasco, and Villas-Boas, Sofia. 2020. "Are We #StayingHome to Flatten the Curve?" May 25, 2020. Department of Agricultural and Resource Economics, first version April 5, 2020. Revised May, 2020. CUDARE Working Papers. <http://dx.doi.org/10.2139/ssrn.3569791>
- Simon, Herbert A. 1954. "Bandwagon and Underdog Effects and the Possibility of Election Predictions". *The Public Opinion Quarterly* 18(3): 245-253.
- St. Onge, Peter, and Gaël Campan. 2020. "The flawed COVID-19 model that locked down Canada. Montreal Economic Institute". June 4. Accessed March 1, 2021. <https://www.iedm.org/the-flawed-covid-19-model-that-locked-down-canada/>.
- Stewart, Heather, and Ian Sample. 2020. "How No 10's relationship with its scientists broke down". *The Guardian*, October 17. Accessed March 1, 2021. <https://www.theguardian.com/world/2020/oct/16/scientists-call-to-arms-went-unheeded-how-the-relationship-with-no-10-broke-down>.
- Walker, Patrick, Charles Whittaker, Oliver Watson, Marc Baguelin, Kylie E C Ainslie, Sangheeta Bhatia, Samir Bhatt, et al. (2020). "Report 12: The global impact of COVID-19 and strategies for mitigation and suppression". *Imperial College COVID Response Team*. March 26. Accessed March 1, 2021. <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-12-global-impact-covid-19/>
- Winsberg, Eric, Jason Brennan, and Chris W. Surprenant. 2020. "How Government Leaders Violated Their Epistemic Duties during the SARS-CoV-2 Crisis." *Kennedy Institute of Ethics Journal* 30(3-4): 215-242.