The Need for a Systematic and Iterated Comparison of Different Policy Models

Mike Bithell[1], Edmund Chattoe-Brown[2] and Bruce Edmonds[3]

[1]U Department of Geography, Cambridge University
mike.bithell@geog.cam.ac.uk

[2] School of Media, Communication and Sociology, University of Leicester
ecb18@leicester.ac.uk

[3]Centre for Policy Modelling, Manchester Metropolitan University
bruce@edmonds.name

Extended abstract prepared for presentation at the 2021 EU Conference on Modelling for Policy support: collaborating across disciplines to tackle key policy challenges

22 – 26 November 2021, online event.

Why Single Models are not Reliable for Informing Policy

In the early 1990s, a moratorium was put on all cod fishing off Newfoundland and Labrador. The cod were declared commercially extinct and around 30,000 people lost their jobs as a result. The Harris Commission's report into the causes of this collapse said that modellers "*…failed to recognize the statistical inadequacies in their bulk biomass model…*" and that they had concerns that "*…weaknesses in scientific management and the peer review process permitted this to happen.*" The scientists and policymakers had become committed to a particular description of reality. As a result, their model was inadequate and this was not picked up. In this case, the policy modelling had contributed to the disaster due to the narrowness of their modelling – it had made things worse (see account in [1]).

Early in 2020, Neil Fergusson and his team used a complex model to simulate the spread of COVID19 under a range of scenarios, forecasting 100,000s of deaths in the UK if policies to reduce its spread were not enacted. This model was *not* the sole basis of the subsequent policy change in the UK, since this was consistent with other mathematical models as well as the unfolding events in Italy. The model was hurriedly adapted from a model developed 13+ years previously concerning influenza. The model was criticized because many people did not like the policy conclusions drawn, but more pertinently due to the fact that the code, "thousands of lines of undocumented C", was not publicly available, and so had not been critiqued and checked by other researchers (see account in [5]).

It is easy for modellers to (a) see the world through their model, developing a myopic view of the world (the effect of "Kuhnian Spectacles") and (b) to not fully understand their own models [4].

Learning from Examples of Model Intercomparison

Thus, the question arises about the reliability of such models for policy purposes. How can policy actors rely on models that they cannot personally understand? Model comparison projects (MIP), such as those in the climate community give some clues (for an account of these see [2]). MIPs have many advantages, including: (a) they allow modellers to build on the past rather than re-inventing the wheel, (b) encourages the independent reproduction and analysis of existing models resulting in their being better understood, (c) help to determine which kinds of models are better for which aspects of problems or what kinds of situation they are applicable in, (d) a continuously updated and refined base of models helps build credibility and (e) form a more credible and robust basis from which to inform policy. Establishing an agreed framework for the exercise and then running MIPs is not easy, and requires sustained effort but can be grown over time, helping to ensure consistency even if individual models come and go. Other domains may not be as ready as that of climate change but MIPs can help a field mature and to provide a more useful and understood tool to inform policy.

Developing Policy Model Intercomparison as Standard Policy Modelling Practice

If there had been model comparison exercises concerning the North Atlantic fisheries – systematically comparing a *variety* of models – the accepted assumptions might have been questioned more and its collapse prevented. If Neil Fergusson had made his code available in 2006 when he published his paper, then the code might have been critiqued and improved by a community of interested researchers over many years. This would have improved the code, making such models more defensible. Such a community of practice might have resulted in reliable models that were: ready for a new pandemic, adapted to be relevant to the policy issues and thus able to help the UK government to react more quickly to events (thus preventing many deaths). As a community of policy modellers, we need to get our act together on important issues – to get out of our bunkers where we are dealing with only single models – to systematically and iteratively compare models, in order to provide a more reliable basis for policy. Reproducing models is the first, most necessary step [3], but for robustness, reliability and depth one needs to compare a *diversity* of models about the same phenomena in a structured manner.

References

[1] Aodha, L. & Edmonds, B. Some pitfalls to beware when applying models to issues of policy relevance. In Edmonds, B. & Meyer, R. (eds.) Simulating Social Complexity - a handbook, 2nd edition. Springer, 801-822 (2017). https://doi.org/10.1007/978-3-319-66948-9_29

[2] Bithell, M. & Edmonds, B. The Systematic Comparison of Agent-Based Policy Models - It's time we got our act together! Review of Artificial Societies and Social Simulation, 11th May 2021 (2020). https://rofasss.org/2021/05/11/SystComp

[3] Chattoe-Brown, & al. Reproduction as a Means of Evaluating Policy Models: A Case Study of a COVID-19 Simulation. medRxiv 01.29.21250743 (2021); https://doi.org/10.1101/2021.01.29.21250743

[4] Edmonds, B. & Hales, D. Replication, Replication and Replication - Some Hard Lessons from Model Alignment. Journal of Artificial Societies and Social Simulation 6(4):11. (2003) http://jasss.soc.surrey.ac.uk/6/4/11.html

[5] Squazzoni, F. & al. 'Computational Models That Matter During a Global Pandemic Outbreak: A Call to Action' Journal of Artificial Societies and Social Simulation 23(2):10 (2020) <http://jasss.soc.surrey.ac.uk/23/2/10.html>.