Logic, Methodology and Philosophy of Science

Proceedings of the Fifteenth International Congress

Hannes Leitgeb
Ilkka Niiniluoto
Päivi Seppälä
Elliott Sober
Editors

20 Modelling failure

USKALI MÄKI *

Abstract. Philosophy of science is largely inclined to portray science as a success story. While not in the least denying the great successes of science, it would seem there is an important and interesting other side of the story, also worth the systematic attention of philosophy of science: developing accounts of the nature, conditions and dynamics of both failure and success should be on the philosophy of science's agenda. In this article I focus on one prominent style of scientific inquiry, that of modelling, and on one part of philosophical study of science, that of offering philosophical accounts of models and modelling in science. A sound philosophical account of modelling should contain resources for identifying and diagnosing modelling failure. The ability to articulate (at least rudiments of) a systematic account of modelling failure can be used as a test of one's account of model and modelling. Here I expose my own account to such a test using an example from economics (its alleged failure in anticipating and conceiving the 2008 financial crisis), showing how my account provides an encompassing framework for identifying and analyzing failures in modelling.

Keywords: models, models in economics, modelling failure, economics, philosophy of economics, financial crisis.

1 Science fails too

Philosophy of science is largely inclined to portray science as a success story. It takes as its task to explain how science manages to acquire truthful information about the world or to generate successful predictions and explanations of phenomena, and it aspires to do this by articulating the procedures and principles of science that make this success possible. While not in the least denying the great successes of science, it would seem there is an important and interesting other side of the story, also worth the systematic attention of philosophy of science – not just as an unfortunate accidental

*University of Helsinki

residuum of failures soon to be corrected, but rather as a massive enduring part of scientific practice with its own regular features. Developing accounts of the nature, conditions and dynamics of both failure and success should be on the philosophy of science's agenda. Ability to produce such accounts should be one of the criteria of success of philosophy of science itself.

As my illustrations here will derive from the discipline of economics, it is interesting to note that there is a striking analogy between conventional economics and conventional philosophy of science. Both are fascinated by the successes of their target domains, the market economy and science, respectively. Consider the following statement by the economist Nouriel Roubini. In promoting an unorthodox approach to the study of the economy ("crisis economics") he stresses the difference between studying failures and studying successes. "Crisis economics is the study of how and why markets fail. Much of mainstream economics, by contrast, is obsessed with showing how and why markets work – and work well." (Roubini & Mihm, 2011, p. 39) My paper is an exercise in the philosophy of science analogue of crisis economics – while much of conventional philosophy of science is analogical to the urge of mainstream economics to show how wonderfully markets function.

Naturally, the ambition to understand failure in science must be divided into manageable portions. Here I focus on one prominent style of scientific inquiry, that of modelling, and on one part of philosophical study of science, that of offering philosophical accounts of models and modelling in science. There are many such accounts available in the literature, and the challenge is to compare them for their credentials. One obvious way to proceed is to check them against empirical evidence concerning actual models and actual modelling practices. And provided we take these practices to include failures, then the capacity of the philosophical accounts in dealing with such failures may be taken as a major criterion of the success of those accounts. A sound philosophical account of modelling should contain resources for identifying and diagnosing modelling failure. The ability to articulate (at least rudiments of) a systematic account of modelling failure can be used as a test of one's account of model and modelling. Here I expose my own account to such a test, showing how it provides an encompassing framework for identifying and analyzing failures in modelling.

2 Failure of model, failure of target: Economics and the financial crisis

'Modelling failure' is ambiguous between *modelling <failure in the target system>* and *failure in <modelling>*. In the first category, the failure lies in the functioning of the target system rather than in an attempt to model the target system. The possibility of such a failure applies to target systems to which we are entitled to ascribe some idea of proper functioning. Scientists can then model such failures of their target systems to function properly, such as heart failure in the human body, business failure,

market failure, failure of materials, failure of machines in manufacturing systems, failure of engineered arrangements such as energy systems and jointed rock slopes to resist landslides, and so on.

Scientists may be successful in modelling failure. Or they may fail in modelling failures in their target systems. Or they may succeed or fail in modelling the proper (non-failing) functioning of their target systems. It is an instance of the second combination that I will use as an illustration in what follows. I will be discussing *tools and acts and strategies of modelling that allegedly fail to model the failures of some target systems* – such as this example from materials physics: "Traditional hyperelastic models of materials ignore the fact that no material can sustain large enough deformations without failure." (Volokh, 2010, p. 684) Here models fail to model failure due to employing the excessively idealizing assumption of hyperelasticity. The illustration I will use later in the paper is of this sort of double failure: failure in economics to model the failures of the financial system. The possible sources of failure, however, will be shown to be far more varied than just a single idealizing assumption.

Just a few years before the outbreak of the financial crisis of 2008 and the subsequent persistent recession, there was a good deal of optimism in the academic air, celebrating the "great moderation" of business cycles and praising macroeconomics for having solved the "problem of depression prevention" by developing theories and methods and policy recommendations that were disputed by few in the economics profession. Here is Nobel Laureate Robert Lucas from Chicago in 2003, in his presidential address to the American Economic Association:

> "The term ['macroeconomics'] then referred to the body of knowledge and expertise that we hoped would prevent the recurrence of that economic disaster. My thesis [...] is that macroeconomics [...] has succeeded: Its central problem of depression prevention has been solved, for all practical purposes, and has in fact been solved for many decades." (Lucas, 2003, p. 1)

Soon thereafter, things went badly wrong (in fact they were getting wrong at the time of Lucas's statement). In a few years, triggered by the subprime mortgage crisis that burst the bubble in the US housing markets, the global financial system would collapse – without anticipations, warnings, or recommendations as to how to prevent it offered by the profession that had just a while earlier congratulated itself for having solved the problem.

The generation of the crisis was no simple process. It is therefore no surprise that the aftermath of the financial crisis of 2008 has exhibited an almost proverbial blame game. Whom, or what, to blame for the crisis? The candidates have ranged from the government (for regulating too little, too much, or wrongly) through the design of the global financial market system (for its inherent instability and susceptibility to systemic risk) to the credit rating agencies (for massive mistakes in their assessments) and, of course, the human nature in general (greed and all that).

The discipline of economics has received its share of the blame. Even those who

charge central bankers like Alan Greenspan and Ben Bernanke for having made fatal mistakes consider that those mistakes were shared by the establishment of the economics profession and so individuals like them cannot bear the main responsibility (e.g. Posner, 2009, p. 286). It is a matter of collective and institutional failure, and this is where the discipline of economics enters the picture, with its collectively held and institutionally ingrained conventions and convictions, principles and practices. Indeed, many people, including many leading economists, started questioning the performance of economics as a scientific discipline, including the explanatory and predictive capacities of the highly abstract mathematical models that have become so popular in the discipline. In July 2009, the cover of *The Economist* asked: "What went wrong with economics?" A bit later, Nobel Laureate Paul Krugman (2009) famously wrote in *The New York Times Magazine*:

> "[. . .] the economics profession went astray because economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth. [. . .] When it comes to the all-too-human problem of recessions and depressions, economists need to abandon the neat but wrong solution of assuming that everyone is rational and markets work perfectly. The vision that emerges as the profession rethinks its foundations may not be all that clear; it certainly won't be neat; but we can hope that it will have the virtue of being at least partly right."

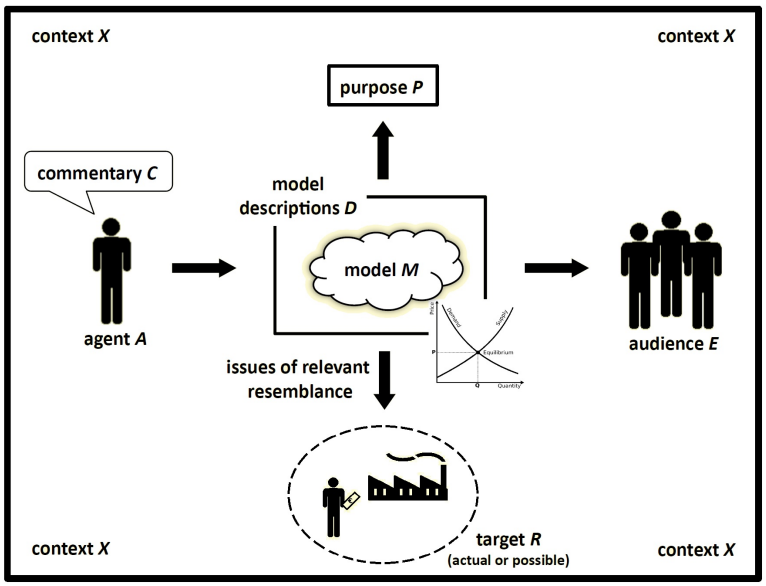
The above passage provides rudiments of a methodological diagnosis of the alleged failures of economics. It claims that economists have been preoccupied with the beauty and neatness of their models, expressed in impressive mathematics, while they have forgotten the task of looking for truths about the real world. As to the contents of their models, the claim is that economists have envisaged a fantasy world of perfectly rational agents in perfectly self-regulating markets, and this fantasy world is too far removed from the imperfections of the real world to be helpful for acquiring truthful information about the latter. The economics profession is criticized for holding an all too strong faith in the powers of the invisible hand, manifesting itself in the widely accepted and applied but allegedly failed DSGE (dynamic stochastic general equilibrium) models in macroeconomics and those of efficient markets in finance.

The remarks in the following sections will provide an elaboration of such rather popular Krugman-style allegations. This will be done by dealing with the alleged failures of economics vis-à-vis the financial crisis as modelling failures and by offering a series of possible partial diagnoses of different kinds and sources of modelling failure. Modelling is a multi-stage and multi-faceted cognitive process, so there are multiple sources of, and multiple opportunities for, possible failure – as well as multiple ideas of what constitutes failure. Some of these will be mapped. For this task, we need an account of model and modeling that is rich enough for exhibiting several such opportunities for failure.

3 An account of model and modelling

In the last couple of decades, philosophy of science has recognized models as among the key cognitive tools in science and modelling as one of the key activities in scientific practice. It has become a major industry within philosophy of science to produce accounts of model and modelling in science. These accounts have often been generated and applied (sometimes even tested) using historical and contemporary case studies in a variety of scientific disciplines, from physics and biology to the social sciences. I will now briefly summarise my own account and then put it into use for highlighting failure rather than success, with economics and the financial crisis serving as the illustrative case. Its capacity in this role provides a test of its general adequacy.

Philosophers of science have moved beyond the idea that models involve a simple two-place representational relationship between the model and its target: *M is a model of target R*. It is now commonplace to conceive of representational models as more complex, involving in addition an agent and a purpose: *Agent A uses (builds, employs) M as a model of target R for purpose P* (see e.g. Giere, 1999). Some philosophers wish to add further components in the modelling relationship, such as some idea of an interpretation (see e.g. Weisberg, 2013). I have proposed that representational modelling is still more complex (e.g. Mäki 2009b, 2011b, 2011a). My account portrays model representation as a multi-faceted activity that can be schematized as [ModRep][1]:



[1]Thanks for brilliant help in designing the figure go to Ilmari Hirvonen

[ModRep] can be put in words only:

> Agent *A*
> uses multi-component object *M* as
> a representative of (actual or possible) target *R*
> for **purpose *P***,
> addressing **audience *E***,
> at least potentially prompting genuine **issues of relevant resemblance** between *M* and *R* to arise;
> describing *M* and drawing inferences about *M* and *R* in terms of one or more **model descriptions *D***;
> applies **commentary *C*** to identify and coordinate the other components;
> and all this takes place within a **context *X***.

Among the novelties here we can list the involvement of an audience; the idea that merely potentially giving rise to genuine issues of relevant resemblance is needed for representation to be in place; a rich functionally loaded notion of commentary; and a generous placeholder for various contextual factors.

A simple version of the idea of *success* in modelling would say that success occurs when relevant resemblance between *M* and *R* becomes secured. This requires, for example, that the causal structure or causal forces captured in the model relevantly resemble those in the target; the model descriptions are of the sort that enable the needed inferences; and that the components of [ModRep] become coordinated and communicated so that purposes *P* and audiences *E* are sufficiently served. Finally, context *C* must be such that it is sufficiently supportive of those accomplishments.

We can then say that *failure* in modelling occurs when relevant resemblance between *M* and *R* fails to be secured, in one way or another, for one or another of the various possible reasons that [ModRep] helps identify. The components [ModRep] and their relations will next be investigated as potential loci and sources of modelling failure. It appears that many existing critiques of economic modelling can be construed as focusing on some specific component in the structure of [ModRep] and that some other possible critiques can also be envisaged within this framework. In both cases, it seems obvious that the claims about possible sources of failure can be made sharper than without the framework of [ModRep] and that the framework will also enable more focused assessments of the credibility of such claims.[2] I will next rush through the framework component by component.

[2] I will use the structure of [ModRep] for outlining *possible* sources of failure without making any strong claims about actual failures. The usefulness of the framework for organising focused scrutiny is not dependent on the correctness of such claims. There is an analogy between the exercise here and much of economic modelling: the latter often describes possible ways in which a pattern of phenomena comes about, and so do the examinations outlined in the present remarks.

4 Multiple sources of modelling failure

I will exhibit the usefulness of the [ModRep] framework for organizing some of the popular misgivings about economic modelling – for identifying and localizing the various criticisms of economic models for their failures regarding the financial crisis (and surely beyond this case as similar criticisms are a chronic feature of debates around economics). It appears that all components listed in [ModRep] have been or can be charged for not having done well.

Agent A

Identifying the agent of modelling as a separate component, and attributing characteristics to it, is somewhat awkward since this is the melting pot where the influences of the other components get together. This is fortified by the fact that individual and collective agency interact and depend on one another. The individuals inventing and proposing models have been educated and socialized in the collective disciplinary culture and prevalent fashions of research practice. On the other hand, the individual modeller must persuade the collective to join her in using an object as a model. The identity and properties of *A* make a difference for whether *A* qualifies as a (credible) economist; what sorts of models are being built and examined, shaped by *A*'s skills and background beliefs; what is taken seriously as a model worth some further attention; and so on.

Economists are generally recognised as intelligent people. Yet the critics argue that this is not sufficient for successful modelling and that the failures regarding the 2008 crisis are one indication of this. They say economists are too narrowly educated (mainly just in contemporary economics, math and statistics), too ignorant about history (of the economy and of their own discipline), about the other social sciences, about culture and human psychology. Some say their competences and epistemic preferences are ill suited for modelling the complexities of social reality. Their mathematically inclined style of inquiry encourages them to streamline the nuances of the real world in epistemically harmful ways: they are extremely skilful in mathematical puzzle solving when reasoning about the model worlds, but relatively speaking clumsy and uninformed in connecting their formulas to the detailed complexities of real world economies. These capacities and their limitations may also nurture overconfidence, hubris, and arrogance – characteristics often attributed to the economics profession and conducive to the sorts of failure witnessed in connection to the 2008 crisis (see e.g. Posner 2009; Fourcade et al. 2015).

Regarding the contents of their worldviews, there is empirical literature suggesting that economists are more self-seeking than other professions, either due to economics education or self-selection (see e.g. Carter & Irons, 1991). This may be suggested to result in systematic biases in favour of models that put too much stress on self-seeking behaviour amongst the populace at large.

. . . uses multi-component object M as

Even though models are systems with multiple components, they are necessarily partial: models isolate only a subset of potentially relevant causal factors in the economy while leaving the rest out of the model. Such isolations – inclusions and exclusions of factors – can be accomplished simply by omitting them without any mention, or by means of idealizing assumptions that assume, directly or indirectly, those factors to be absent, constant, or otherwise ineffective (Mäki, 1992).

An important potential source of modelling failure is, of course, the contents of the model system itself. Failure occurs if causally, explanatorily and/or predictively relevant factors are excluded from the model. Virtually all criticisms of the performance of economics regarding the 2008 crisis make this charge. Here is Nobel Laureate Joseph Stiglitz: "the conventional models inadequately modelled – and typically left out – many, if not most, of the key factors that played a central role in this crisis" (Stiglitz, 2011, p. 172).

Models of efficient financial markets rely on strong idealizing assumptions such as zero transaction costs and perfect and symmetrical information between the agents. These idealizations help produce an image of the financial system in which market prices fully incorporate all relevant information and in which there can be no bubbles in asset prices such as those of stocks or houses. This is a model system that has the features of being self-regulated and having the capacity of containing all relevant risks without external regulation (so Alan Greenspan and others believing in these models were able to relax). Critics argue that such an exclusion of real-world imperfections from the models is fatal, suggesting that the properties of real-world markets may in fact be the reverse of those of the model-world markets: "... where the Efficient Markets Hypothesis suggests that financial markets provide a way of managing economic risk, the evidence suggests that they are actually a major source of risk." (Quiggin, 2010, p. 51)

Macroeconomic models, in their turn, have employed the highly idealized notion of representative agents. These models have chronically missed the crucial causal factors such as animal spirits, herd behaviour, informational asymmetries, structure of financial markets, and corporate governance, therefore failing to recognize phenomena such as excess indebtedness, debt restructuring, bankruptcy, and agency problems. The critical verdict is that any model with these characteristics "leaves out much, if not most, of what is to be explained; if that model were correct, the phenomena – the major recessions, depressions and crises that we seek to understand – would not and could not have occurred" (Stiglitz 2011, 168; see Akerlof and Shiller 2009). So the failure regarding this component is a matter of leaving out from the models important (e.g. bubble-generating) factors and mechanisms that are responsible for the sort of crisis we have just experienced.

... a representative of (actual or possible) target R

The key point of modelling is that models stand for their targets as their representatives, and that models are directly examined for their properties and behaviour so as to indirectly acquire information about their targets. This involves a sort of objectifi-

cation or reification of model systems as such direct subjects of inquiry. One simple failure lured by this is the failure to distinguish between the model and the target in one's reasoning, to proceed, without further systematic inquiry, as if the model system is the target system, and to believe that the properties discovered to be possessed by a model are also properties of its targets. This is a very old criticism of economic models, levelled already in the 19th century. It is still topical today, also in relation to the modelling failures regarding the 2008 crisis.

The targets of models can be either actual or merely possible objects or systems, and whether they are one or the other has consequences for what kind of information modelling can deliver, and what sorts of failure may occur. Even though models can be required to ultimately stand for some actual target objects or systems (such as really existing macro economies or the current global financial system), economists often treat their models as representatives of some possible targets, using them as tools for how-possibly explanations. Often an economist simply takes on the task of examining whether it is possible to derive a given stylized fact from the assumptions of individual optimizing behaviour, that is, whether it is conceivable that the stylized fact is an outcome of the functioning of a mechanism involving such behaviour. Or an economist may examine the conditions under which a simple imagined market system is stable. It may happen that no one (or no contemporary economist, or no sufficient number of economists credible enough to have their voices heard) takes the task of establishing whether these connections also obtain in some actual target system. This means that no information is generated about actual systems. Such a situation may miss important facts about the relevant actual target systems, such as the global financial system.

One can conceive of the very idea of a model's target with different degrees of specificity, in terms of a varying magnitude of attributes. With a sufficiently rich notion of target, one can then argue that the failure of economic modelling was to be preoccupied with relatively stable, bubble-free economies as the main or sole targets, and this meant missing the actually emerging dramatic instabilities in the bubble-generating crisis-prone economies. This is to say those models deal with "special cases where market inefficiencies do not arise" (Stiglitz, 2011, p. 166) or that they do not apply to economies that are capable of generating bubbles. In other words, the targets of the criticized models were too narrowly conceived, especially as they crowded out actual targets that mattered greatly at the time. Ben Bernanke (2010) acknowledges the premise of this reasoning, but uses it for defending macroeconomic models, arguing that they work fine under ordinary conditions: "Economic models are useful only in the context for which they are designed. Most of the time, including during recessions, serious financial instability is not an issue. The standard models were designed for these non-crisis periods, and they have proven quite useful in that context." (2010, p. 17) In other words, the 2008 crisis was not a feature of the proper targets of the then-dominant economic models. Thus the problem was not with those models but with the modelling practices that had failed to address targets with bubble-generating capacities.

. . . for purpose P,

There is a vast multiplicity of possible purposes that models – or any given model – can serve, or intended or presumed to serve. These include explaining this or that phenomenon or aspect of the phenomenon; predicting this or that phenomenon with this or that accuracy of timing, magnitude etc.; exploring the properties of possible scenarios; solving this or that puzzle (so as to get yet another paper published); elaborating this or that new or old technical tool; designing an institution or policy with these or those desired features; educating or persuading students (or lay people, media, politicians); and so on.

It makes no sense to talk about successes or failures without specifying the relevant purposes. Success and failure are functions of purpose, among other things. This provides part of the pragmatic context of models. Modelling can serve multiple purposes, such that any particular model is designed for – or is expected or found to serve - one or more particular purposes and can only be evaluated in terms of its success in serving those purposes. It is worth noting that the issue of target specification or the proper domain of a model's applicability (that we just briefly discussed above) can often be transformed into an issue of the purposes of modelling. This is what Stiglitz suggests: "Is the purpose of an economic model to help us predict a little bit better how the economy is performing in 'normal' times – when things do not matter much? Or, is the purpose of an economic model to predict, prevent and manage big fluctuations and crises?" (Stiglitz, 2011, p. 168)

Stiglitz and many others charged economics for *predictive failure*, but this is very ambiguous. Of the 2008 financial crisis, one may have failed to predict (a) *that* it is coming; (b) *by what pathways* it is coming; (c) *when* it is coming; and (d) *with what (e.g. quantitative) characteristics* it is coming. These tasks are concerned with the anticipation of *actual* events, and they are not equally easy. While (a) is surely the easiest task, (c) and (d) are far more difficult – while their degree of difficulty depends on the desired accuracy of the predictions. What is important is that many economists defended their models and modelling practices by arguing that prediction is not among their proper and reasonable goals after all.

An even heavier charge is for a *failure of conception*: economists not only failed to predict the crisis, they failed to conceive it, which is even more serious. As much as economists are fond of modelling possibilities, it is notable that the sort of systemic crisis exhibited by the financial system in 2008 was not among the possible scenarios entertained by economists. They failed to conceive what is *possible* within the system given the sorts of causal mechanism that the system contains. And they failed to conceive what the system has a *propensity* to do, given those mechanisms.

Many critics have pointed out that the dominant purposes of large and most respectable parts of economic modelling may have been removed too far from the timely needs of effective economic policy – such as the anticipation of certain dramatic kinds of real world development or even the recognition of their possibility based on accounts of the complex system of underlying mechanisms. The suspicion is that economists are in-

clined to address relatively small and easy specialized intra-academic puzzle-solving challenges so as to maximize their academic performance measured in terms of publication output, rather than to tackle big and difficult issues that take more time and effort and are more risky in their capacity to yield an impressive publication record. Perhaps the dominant purposes of modelling have become rather fragmented and inward looking.

These observations inspire two further remarks. The first is that given the large variety of possible purposes, there is also a large variety of ways of failing (so it is easy to fail) as well as of ways of escaping charges of failure (so it is difficult to fail). One can defend a model or a model format by varying the associated purposes. In response to (alleged) failure in relation to (intended or presumed) purpose P_1 the modeller makes an excuse or escape by proposing an allegedly more appropriate purpose P_2 and by claiming success in relation to P_2 (and in response to the possible further failure in relation to P_2 the modeller escapes by proposing another purpose P_3, and so on). No doubt this structure has been in place in the debates around economics. The second remark is that there is a limit to the room that can be permitted for escape. The purposes of modelling themselves are and should be subject to critical scrutiny and debate. Not all possible purposes are equally appropriate. Some purposes just cannot be ignored without damaging the legitimacy of a modelling discipline.

. . . addressing audience E,

Modelling is always addressed to some audience or set of audiences. Models are used to convey information to audiences, to educate them, to impress them, to persuade them, to use them as test partners. The possible audiences include likeminded experts in the modeller's specialised field of inquiry and those in disagreement, other members of one's own discipline and those in other disciplines, students and journal editors, media and policy makers, and so on. The expectations and beliefs of particular target audiences partly determine what is offered to them by the modellers. This provides another pragmatic aspect of modelling. (One can of course try to embed the roles of audience in the category of purpose, but this would make the latter intractably extensive and variegated and internally structured, so I prefer keeping audience as a separate component.)

I mentioned using an audience as a test partner. Any model, in the course of its lifetime, travels through a private-public dimension, from private conception to public acceptance or rejection. At various stages, audiences as test partners can be internal (modeller herself) or external (other scientists or non-scientists). The internal audience as a test partner examines the conceived model for its public recognisability and acceptability, at least for further study and discussion, before submitting it to public scrutiny. Anticipations of the forthcoming public reception of the model are shaped by the dominant modelling conventions in the discipline or research field, so the model had better be constructed and proposed without violating these social conventions to have a chance. The public examination of models will therefore virtually always restrict itself to a limited and relatively homogeneous set of candidate models (see Mäki, 1993, p. 97). This is one mechanism resulting in the sort of intellectual herd behaviour

that has been mentioned among the sources of modelling failure regarding the 2008 crisis.

Among the external audiences, there is the primary audience consisting of other academic economists attending specialised seminars and reading top journals in the same field. One of the important functions of this audience is to serve as a pool of referees used for assessing paper submissions for publication in journals. Given how academic performance is nowadays measured, this is crucial for academic survival and promotion, so modellers had better be successful in persuading this audience. This takes place within a narrow specialism, which means that critical or corrective feedback tends to be limited, focused mostly on technical details. Broader issues and those of societal relevance play side roles at most. The big picture easily gets lost.

Then there are secondary audiences, such as economists and non-economists in other fields, and policy makers and the general public as well as journalists mediating between academic research and these non-academic audiences. The demand for the big picture may be pressing, but the demand is not very likely to be met due to the overriding pressures of specialised academic performance. In general, there tends to be little effective feedback on the relevant models from these secondary audiences, and this may be due to a lack of competence (e.g. insufficient mathematical literacy), or perhaps in some cases ideological coherence or political convenience.

So it is possible that much of the most respectable academic work in economic modelling has been directed predominantly to other academic economists within some limited field of specialization rather than to audiences busy with policy work concerned with concrete complexities of real world issues. One consequence of this may have been a failure to develop sound comprehensive accounts of the functioning (or malfunctioning) of the system as a whole. Or, in case economic research has been addressed to people in policy practice, the latter may have been happy to receive them without complaints because the economy was believed to be running its course sufficiently smoothly and perhaps because the message from the academia was in line with their political prejudices. Indeed, one of the suspicions presented by the critics is that the belief in the self-stabilizing capacities of unregulated markets has been strong both among leading economists and among leading politicians, and that these beliefs have fortified one another (see e.g. Stiglitz 2010 and Quiggin 2010).

. . . at least potentially prompting genuine issues of relevant resemblance between M and R to arise;

While many different things can stand for, or can be used as a representative of, a given target thing, only those that resemble the target in relevant ways are useful for generating true information about the target. To find out whether this is so, the issue of relevant resemblance must arise. The notion of relevant resemblance involves both objectively ontological and subjectively pragmatic aspects. Resemblance is a matter of objective matter of fact, while relevance is a matter of relativity to some purposes, goals, interests, and audiences. One and the same resemblance relation between a

model and its target may be sufficiently relevant for some purpose and audience but not so for another. The critic may then argue that the dominant available models do not resemble their targets relevantly for conceiving, anticipating, and avoiding the sorts of crisis that was witnessed in 2008. It is important to see that there are two ways in which models may fail to relevantly resemble their targets (Mäki, 2009a, 2011a).

One kind of failure occurs when relevant resemblance is pursued by the modeller, but the attempt fails. In this case *the issue of relevant resemblance arises*, but it will be resolved in support of the conclusion that a given model does not relevantly resemble some target. Or, it may be that this would be the right conclusion, but it is not actually drawn by economists for one reason or another. In either case, *surrogate modelling* is exercised: models are conceived as *bridges* to their targets, so that finding about the properties of models are intended and hoped to be informative about the respective properties of the targets. But the bridge may fail, and the intended resemblance is not established. The model may be based on isolations that serve to exclude, rather than include, some important crisis-generating mechanisms in operation in the real economic system. This would be one type of failure of resemblance, based on *trying and failing*.

Another kind of failure occurs when there is *no trying at all*, unlike in the case of surrogate modelling. It is not a failed attempt but rather a failure to attempt. In what I call *substitute modelling*, models are conceived as *islands* (rather than bridges), with *no issue of relevant resemblance being prompted at all*. Imaginary model worlds are examined for their own sake, as it were, with economists learning a great deal about the properties of their models but nothing about any real world targets – not even about what these targets are not like. The study of models substitutes for the investigation into anything beyond them. Criteria other than those pointing towards resemblance between models and their targets dominate the exercise. This is probably what Paul Krugman had in mind when writing, "economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth" (see also Hodgson, 2009). Krugman is not alone with this suspicion.

The distinction between surrogate modelling and substitute modelling is not crystal clear though, and this creates some room for justifying modelling practices that do not immediately address issues of resemblance. As I said before, it is part of the nature of modelling that attention and research effort are directly focused on examining the properties of models. There are two dimensions along which to consider the further issue of whether such attention and effort are a matter of surrogate or substitute modelling. One dimension is that of *individual-collective*: the task of addressing the issue of resemblance is delegated to other researchers within the scientific community in line with some principle of division of scientific labour. Even if one modeller may appear to practice substitute modelling only, this is supplemented by others who turn the exercise into one of surrogate modelling. Another dimension is the historical one of *now-later*: there is an appropriate temporal order of research tasks, and the task of addressing issues of resemblance will be taken up and accomplished by

some later (generations of) researchers. Both of these dimensions have been utilised by economists in defending their models. (See Mäki 2009a, 2011a.)

The critical diagnosis, however, is to argue that major parts of economic modelling may have degenerated into the investigation of imaginary model worlds only, with no issues of resemblance being raised under the prevalent division of scientific labour or along a natural temporal sequence of intellectual effort. This situation results in treating models as substitute systems that are investigated in their own right with no concern whatsoever with how they might connect with real-world systems. Such a tendency may have been supported by a variety of factors, such as the increasing specialization of economic inquiry, its over-mathematization, and perhaps in some cases by the fact that economists and some of their audiences have been charmed by the smooth functioning of model economies within the model worlds, apparently justifying non-interventionist attitudes in policy making.

. . . describing M and drawing inferences about M and R in terms of one or more model descriptions D;

A given model can be described variously, such as in terms of verbal means, algebra, geometry, diagrams, etc. It is in terms of whatever medium is used for describing the model world that inferences are drawn about the properties and behaviour of that model world. Each type of medium has consequences for many other things in the modelling exercise, such as the range of inferences and claims that can be made (and are likely to be made) about the models, the range of ingredients to be included in the models, and the sorts of audience that can be reached.

In economics education, addressing students from introductory to advanced levels, one and the same model can be described variously at different stages of the education process, beginning with using easily accessible mathematics and gradually employing ever more demanding techniques. In policy memos, one may explain the key message of a model verbally, perhaps using diagrams, and then adding an appendix with a full-blown mathematical description of the model (with goals such as exposing the model to expert scrutiny or impressing the policy makers with a façade of scientificity). The structure of journal articles in terms of model description techniques also resonates with issues of audience, such as the readability of published articles and the sales and citations attracted by the journal.

It has been argued that the popularity of sophisticated mathematical means of description may have fortified the allegedly great distance between some economic models and the real world. This charge can be made more precise by invoking the idea of formal tractability as a dominant guiding principle of modelling (see e.g. Hindriks, 2006). The claim then would be that economists may have been excessively and uncritically constrained and inspired by considerations of formal tractability rather than empirical adequacy or relevant resemblance, and that this has imposed serious limitations on how they view the world. The discipline suffers from "the obsession with technique over substance" (Hodgson, 2009, p. 1210).

Recall Krugman's claim that the charm of "impressive-looking mathematics" has

taken an upper hand and suppressed an interest in truth about the complexities of real world economies. In his vigorous response to Krugman, Chicago economist John Cochrane Cochrane (2011, p. 39) acknowledges that real world economies are more complex than the models economists build and use, and that it would indeed be desirable to be able to model the frictions and other imperfections in the world – and that, unfortunately, the mathematical tools now available are limited for the purpose: "Frictions are just hard with the mathematical tools we have now." So it seems Cochrane admits that economic models have missed important features of real world targets and that part of the reason lies in mathematical descriptions of the models constraining the contents of those models. But he thinks the problem is not with mathematics per se but rather with the limitations of the sorts of mathematics now used: "The problem is that we do not have enough mathematics."

. . . and applies commentary C to identify and coordinate the other components

By themselves, model objects do nothing to serve a purpose, to reach an audience, or to link with a target object. Model objects themselves and their relations to such other things are complex and typically not fully transparent, so they require clarification and coordination. This is provided by the model commentary. Yet there is no guarantee that modellers always have the sort of complete understanding of the complexity of the modelling exercise that would enable them to provide a commentary that successfully spells out what is not transparent. Economists busy with mere substitute modelling might even not have a strong interest in providing a clarifying commentary accurate and comprehensive enough to deal with intricate connections with the real world.

A major challenge is to be clear about the roles of and interrelations between what easily appear as outrageously false idealizing assumptions used in describing theoretical models. Commentary has the important task of accomplishing a functional decomposition amongst model components (e.g. Mäki, 2011c). Idealizations are not to be interpreted literally at face value, they typically need first to be translated into claims whose roles are well understood in the structure of the modelling process, so that they can be more directly assessed for how well they serve their proper functions. These latter claims may turn out to be true statements about things such as negligibility (of some falsehood for some purpose) or applicability (of the model to one rather than another domain), or they may become justified as serving a useful purpose as early steps in a sequence of models leading to ever better ones. Or they may fail in playing their proper roles. To be able to check whether they fail or succeed, an informed commentary must have done its job of functional decomposition. Economists often fail to provide such a commentary.

Model commentary also resonates with audiences. Given the reputation of arrogance and self-confidence enjoyed by the economics profession, it is remarkable that when addressing *fellow economists*, they exhibit relatively more modesty and humility when presenting their models and making claims about them. They pay more attention to the complexity of the issues, and are more explicit about the various provisos, background assumptions, and uncertainties that are involved. When addressing *non-economists*

on the other hand, there is less modesty and humility, and less attention is paid to the complexities hidden in the background. "In private discussions among ourselves we recognize this complexity, but we don't add the appropriate warning labels to our models when they are discussed in public. There, we pretend we understand more than we do." (Colander, 2011, p. 20) Another version of the story, often told by economists, is that their otherwise sound ideas are misinterpreted and ill applied by policy makers due to incompetence or the involvement of messy conflicts of political interest. This charge may be intended to redirect the blame for failure, but it can also be taken to emphasise the high importance of supplementing (sound or unsound) models with a sound commentary, but it is not obvious that economists are doing their best in this regard.

It should be part of the academic competences of trained economists to be able to be clear about what their models are for; what the models are about; what the models are capable of doing, and what not; how reliable the models are; what sorts of criticisms have been levelled against the models and how the criticisms have been responded; what alternative models there are; etc. The challenge is not easy, and it is clear that it has not been met with sufficient exuberance and success. The capacity of writing "warning labels" would be part of the needed professional competence. Colander suggests requiring "publications of models that seem to have policy relevance to include an explicit warning directed at the non-scientific users of the model" (Colander, 2010, p. 424). Such warning labels would alert the relevant audiences to the capabilities and limitations of the models.

It is clear that the commentary attached to popular economic models has failed in spelling out their potentials and limitations. This has led to existing models being used for inappropriate purposes and non-existing but important models not having been built at all. Existing models depict bubble-free economies, but they have been erroneously applied to the current financialized economies that are not bubble-free; and models that would be adequate for representing our bubble-prone economies have failed to be developed by a sufficient number of sufficiently credible economists. Avoidance of these errors would be helped by reliable commentaries informing what any given model can and cannot do. It should be added here that philosophers, methodologists and historians of economics might be expected to be well equipped to do their share in contributing to the construction of an enlightening commentary of economic models.

. . . and all this takes place within a context X

The final component of the context of modelling includes lots of various further ingredients that make a difference for models and modelling practices. It includes items such as intra-disciplinary conventions and practices, standards and incentives, arrangements of education, research and publishing, and so on. The context also includes various external (including non-academic) enabling and constraining conditions, expectations, pressures, resources, such as the ongoing transformation of the university institution and the societal status of economics. Such contextual factors make a direct or indirect impact on the other components of the system of modelling.

For example, the deficiencies in the competence of the economics profession to provide adequate commentary of the capabilities and limitations of the models it produces is likely to be an outcome of the rather narrow education nowadays offered to economics students – consisting mainly of recent economic theory and mathematical and statistical techniques. It is generally recognized that fatal consequences may follow from not systematically learning about the other social sciences, about the society at large, about past economic history, about the history and philosophy of the economics discipline itself.

Thinking of the allegedly flawed contents of the models behind the 2008 crisis more directly, the structure of the disciplinary division of intellectual labour may have played important roles. That the isolations of macroeconomic models have excluded causally relevant details of financial markets is partly a product of the structure of disciplinary practices within economics, namely the increased specialization that has deepened the divide between the two fields: "The fact that finance and macroeconomics have become separate fields with some difficulties of intercommunication may have been the inevitable result of the relentless pressure for ever-greater specialization in academic disciplines." (Posner, 2009, p. 328)

As to the epistemic values and conventions underlying disciplinary practices, the ideal of formal tractability may have favoured certain problematic idealizations and suppressed the pursuit of factual adequacy. A tendency towards substitute modelling may have been supported by failures in the division of labour in economic inquiry (economists supposedly bridging the gap between theoretical models and their targets not showing up); by the compulsory incentive to publish elegant modelling exercises in prestigious journal with little need to address issues of real-world connections; and perhaps in some cases by economists and some of their non-academic audiences being overly charmed by the smooth functioning of model economies, therefore reluctant to consider the importance of any real-world imperfections.

Exceptional amongst the social sciences is the role of the economics discipline in contemporary society, the intellectual and political authority economics enjoys regardless of its failures. Above, I cited Colander's confession, "we pretend we understand more than we do" and we could add that economists do so in order to – or with the consequence of – protecting and promoting their socially acknowledged authority. In the worst case, there is a nightmarish scenario on which the more economists are consulted for policy advice, the more they need to pretend to know, and so the higher the likelihood of policies going astray. Avoiding the nightmare would require some smart restructuring of the institutions of the economics discipline.

***

Models and modelling as such are powerful means of acquiring information about dynamically complex systems such as economies. While powerful, they are also prone to error and epistemic risk – and not just epistemic risk, but institutional risk as well in that the (academic and other) institutions of economic inquiry may fail to provide appropriate incentives and other preconditions for adequate modelling. Investigating

the possibility and actuality of modelling failure is a matter of exercising disciplinary risk management (see Mäki, 2011a).

Bibliography

Akerlof, G., & Shiller, R. J. (2009). *Animal spirits. How human psychology drives the economy, and why it matters for global capitalism*. Princeton: Princeton University Press.

Bernanke, B. (2010). Implications of the financial crisis for economics. Remarks at a conference at Princeton University, 24.10.2010.

Carter, J. R., & Irons, M. D. (1991). Are economists different, and if so, why? *Journal of Economic Perspectives*, *5*, 171–177.

Cochrane, J. H. (2011). How did Paul Krugman get it so wrong? *IEA Economic Affairs*, 36–40.

Colander, D. (2010). The economics profession, the financial crisis, and method. *Journal of Economic Methodology*, *17*, 419–427.

Colander, D. (2011). How economists got it wrong: A nuanced account. *Critical Review*, *23*, 1–27.

Fourcade, M., Etienne, O., & Yann, A. (2015). The superiority of economists. *Journal of Economic Perspectives*, *29*, 89–114.

Giere, R. (1999). *Science without laws*. Chicago: University of Chicago Press.

Hindriks, F. (2006). Tractability assumptions and the Musgrave-Mäki typology. *Journal of Economic Methodology*, *13*, 401–423.

Hodgson, G. M. (2009). The great crash of 2008 and the reform of economics. *Cambridge Journal of Economics*, *33*, 1205–1221.

Krugman, P. (2009). How did economists get it so wrong? *The New Your Times Magazine*, *2 September*.

Lucas, R. E. JR. (2003). Macroeconomic priorities. *American Economic Review*, *93*(1), 1–14.

Mäki, U. (1992). On the method of isolation in economics. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, *26*, 319–354.

Mäki, U. (1993). Social theories of science and the fate of institutionalism in economics. In B. Mäki U. Gustafsson, & C. Knudsen (Eds.), *Rationality, institutions and economic methodology* (pp. 76–109). London: Routledge.

Mäki, U. (2009a). MISSing the world: Models as isolations and credible surrogate systems. *Erkenntnis*, *70*, 29–43.

Mäki, U. (2009b). Realistic realism about unrealistic models. In H. Kincaid, & D. Ross (Eds.), *Oxford handbook of the philosophy of economics* (pp. 68–98). Oxford: Oxford University Press.

Mäki, U. (2011a). Contested modelling: The case of economics. In *Models, simulations, and the reduction of complexity*. Hamburg Academy of Sciences.

Mäki, U. (2011b). Models and the locus of their truth. *Synthese*, *180*, 47–63.

Mäki, U. (2011c). The truth of false idealizations in modelling. In P. Humphreys, & C. Imbert (Eds.), *Models, simulations, and representation* (pp. 216–233). London: Routledge.

Posner, R. A. (2009). *A failure of capitalism. The crisis of '08 and the descent into depression*. Cambridge, MA: Harvard University Press.

Quiggin, J. (2010). *Zombie economics. How dead ideas still walk among us*. Princeton: Princeton University Press.

Roubini, N., & Mihm, S. (2011). *Crisis economics. A crash course in the future of finance*. London etc: Penguin Books.

Stiglitz, J. E. (2010). *Freefall. Free markets and the global economy*. London etc: Penguin Books.

Stiglitz, J. E. (2011). Rethinking macroeconomics: What went wrong and how to fix it. *Global Policy*, *2*, 165–175.

Volokh, K. Y. (2010). On modeling failure of rubber-like materials. *Mechanics Research Communications*, *37*(8), 684–689.

Weisberg, M. (2013). *Simulation and similarity. Using models to understand the world*. Oxford: Oxford University Press.

Author biography. Uskali Mäki: Academy Professor, Academy of Finland; Professor of Practical Philosophy, University of Helsinki; Director of the Academy of Finland Centre of Excellence in the Philosophy of the Social Sciences; Former Professor and Director at Erasmus Institute for Philosophy and Economics [EIPE], Erasmus University of Rotterdam (1995-2006); Former Editor of Journal of Economic Methodology [JEM] (1996-2005); Former Chair of the International Network for Economic Method [INEM] (2007-2008).