*Brit. J. Phil. Sci.* **61** (2010), 541–567

# Economic Modelling as Robustness Analysis

Jaakko Kuorikoski, Aki Lehtinen
and Caterina Marchionni

## ABSTRACT

We claim that the process of theoretical model refinement in economics is best characterised as robustness analysis: the systematic examination of the robustness of modelling results with respect to particular modelling assumptions. We argue that this practise has epistemic value by extending William Wimsatt's account of robustness analysis as triangulation via independent means of determination. For economists robustness analysis is a crucial methodological strategy because their models are often based on idealisations and abstractions, and it is usually difficult to tell which idealisations are truly harmful.

## 1 Introduction

Modern theoretical economics largely consists of building and examining abstract mathematical models. A substantial portion of this modelling activity is devoted to deriving known results from alternative or sparser modelling assumptions. Why do economists spend so much time and effort in deriving the same results from slightly different assumptions? The key to understanding this practise is, we propose, to view it as a form of robustness analysis, in other words as the systematic examination of the robustness of modelling results with respect to particular modelling assumptions.

doi:10.1093/bjps/axp049

Robustness analysis was first explicitly identified as an important strategy for analytic model building by the biologist Richard Levins, but we argue that similar considerations give it significance in economics as well. Surprisingly, philosophers of economics have only recently become interested in robustness. In a recent paper Woodward ([2006]) correctly observes that in economics it is typically regarded as a 'Good Thing'. He also points to the fact that there are different kinds of robustness and that arguments in favour of one kind do not straightforwardly carry over to the others.[1] In this paper we are concerned with only one type, what Woodward calls 'derivational robustness', in other words the robustness of a given theoretical result with respect to different modelling assumptions. Consequently, our focus is limited to theoretical economic models.

Woodward claims that derivational robustness does not provide any additional epistemic credence to the conclusion (see also Cartwright [1991]; Sugden [2000]). If he were right, a significant portion of theoretical model building in economics would have no epistemic value. We take issue with this position, deploying William Wimsatt's ([1981]) account of robustness analysis as triangulation via independent means of determination (see also Weisberg [2006a]). Fairly varied processes or activities such as measurement, observation, experimentation, and mathematical derivation count as forms of determination. Triangulation may involve more than one of these forms (e.g. when the same result is obtained by experimentation, derivation, and measurement) or concern only one of them: the same result can be obtained by different experiments or, as in our case, by different theoretical models. Defined this way, robustness is an epistemic concept in that it is a property of the representational means in our epistemic practises (such as modelling or inference) rather than an attribute of the system being investigated.

The aim of robustness analysis is to distinguish 'the real from the illusory; the reliable from the unreliable; the objective from the subjective; the object of focus from the artefacts of perspective' (Wimsatt [1981], p. 128). For derivational robustness to count as a form of triangulation via independent means of determination, the different derivations of the same result should be somehow independent. But the different theoretical models used to assess the robustness of a result usually share many assumptions. Our claim is that independence of a modelling *result* with respect to particular modelling assumptions may nonetheless carry epistemic weight by providing evidence

---

[1] The argument also works in the converse direction: arguments *against* one kind of robustness do not necessarily apply to other kinds. For example, we agree with Hoover and Perez's ([2000], [2004]) claim that there is no reason to expect that a true econometric model yields results that are robust with respect to whether or not particular variables are included in the model-specification. However, their argument concerns what Woodward calls 'inferential robustness' but an analogous claim does not seem to apply to derivational robustness.

that the result is not an artefact of particular idealising assumptions. In particular, we argue that although robustness analysis is not an empirical confirmation procedure in any straightforward sense, its epistemic value stems from two distinct but intertwined functions. First, it guards against error by showing that the conclusions do not depend on particular falsehoods. Secondly, it confirms claims about the relative importance of various components of the model by identifying which ones are really crucial to the conclusions (cf. Weisberg [2006a]).

The two-fold function of derivational robustness analysis is important in economics for the following reasons. First, it is difficult to subject economic models to conclusive empirical tests. Secondly, economic theory does not always indicate which idealisations are truly fatal or crucial for the modelling result and which are not. Finally, theoretical economic models are always based on various idealisations and abstractions that make some of their assumptions unrealistic (Wimsatt [1987]; Mäki [1992], [1994a], [1994b], [2000]; Weisberg [2006a]).

Since there are no natural constants or numerically exact laws in economics, it may not be possible to measure how 'far' from the truth any given assumption is. Furthermore, even if we knew how far from the truth a given assumption was, such knowledge would often be irrelevant for our epistemic and pragmatic purposes (Melitz [1965]): what really interests us is whether a particular deviation from the truth matters for a given result or not. Robustness analysis helps in assessing this question by providing information on whether or not the particular falsity exhibited by some assumption is responsible for a given modelling result.[2]

This paper is organised as follows. Section 2 provides an introductory discussion on the notion of robustness in a review of the existing literature. Section 3 sheds light on the practise of robustness analysis in economics. Section 4 addresses the criticism that robustness is a non-empirical form of confirmation. To illustrate our claims regarding robustness analysis and its two-fold function, in Section 5 we present a case study, geographical economics. We focus on those characteristics that are representative of the way in which robustness analysis proceeds in economics. Addressing the criticisms levelled against robustness analysis, we discuss the independence of tractability assumptions in Section 6. Section 7 provides the conclusion of this paper.

---

[2] The integration of different kinds of empirical data with theoretical models could also be approached from a general robustness analysis perspective. Here however we are only concerned with derivational robustness. This should not be taken as supporting the attitude of some economists who place excessive weight on analytic solvability (Lehtinen and Kuorikoski [2007a]) at the expense of the integration of theoretical models with empirical data and with other sciences.

## 2 Making Sense of Robustness

As Wimsatt uses the term, robustness means the stability of a result under different and independent *forms of determination*. It provides epistemic support via triangulation: a result is more likely to be real or reliable if a number of different and mutually independent routes lead to the same conclusion. It would be a remarkable coincidence if separate and independent forms of determination yielded the same conclusion if the conclusion did not correspond to something real. The meaning of triangulation here is similar to that in social-science methodology: the use of different methods for cross-checking results. If a feature or pattern is discernible from multiple different perspectives, it is unlikely that it is an artefact of a particular perspective.

Experimental triangulation represents a major instance of robustness analysis. By performing experiments that rely on different techniques and background theories, scientists can make sure that a putative phenomenon of interest (as in Bogen and Woodward [1988]) is not merely an artefact of a particular experimental set-up. Similarly, if different measurement modes produce coherent results, this provides evidence for the existence of a single property that is being measured. In the philosophy of science, the multiple determination of Avogadro's constant provides the most celebrated example of such an 'argument from coincidence'. The experimental and measurement robustness of Avogadro's constant is taken to provide irrefutable grounds for the reality of molecules (Hacking [1983], pp. 54–5; Salmon [1984], pp. 214–20).

Every experimental set-up and means of measurement has its errors and biases. Sometimes we have prior knowledge of these problems, but an element of residual uncertainty concerning the validity of an experiment or a measurement always remains. *Independent ways of determining the same result reduce the probability of error due to mistakes and biases in those different ways of arriving at the result.* Wimsatt generalises this principle to all forms of fallible ampliative inference. Fallible thinkers are better off avoiding long inference chains because the chain as a whole is always weaker than its weakest link. In contrast, in the case of multiple and independent forms of determination, the end-result is more secure than even the strongest individual reasoning. For Wimsatt, all procedures of using various types of robustness considerations in order to distinguish the real from the artefactual count as robustness analysis, regardless of whether there are one or more types of means of determination involved (e.g. laboratory experiment, field experiment, and statistics) (Wimsatt [1987]).

The requirement of independence between means of determination appears evident from this error-avoidance perspective. For two or more of them to provide epistemic security in the form of robustness, they should not share the same errors and biases in the light of prior knowledge. If a given method of determination is independent of another, the probability of its failing to achieve the correct result should not depend on whether the other fails. Independence of errors therefore means that *given that the result holds* (or that it does not hold), the success of a particular method of determination in arriving at the result is independent of whether the other methods reach the correct conclusion.[3] If the methods are independent in this sense and more reliable than pure chance, it is easy to show that observing multiple independent results should increase our belief in the result (Bovens and Hartmann [2003], pp. 96–7). In the following sections we generalise this principle from experiments and measurements to theoretical modelling. In effect, we treat theoretical models as forms of determination.

## 3 Robustness in Economics

According to Michael Weisberg's ([2006a]) account, robustness analysis of theoretical models includes four 'steps': (i) determining whether a set of models implies a common result, $R$; (ii) analysing whether this set has a common structure, $C$; (iii) formulating the robust theorem that connects the result $R$ to the common structure $C$; and (iv) conducting a stability analysis to see whether the connection between the common structure and the result is robust with respect to particular parameter values. The way in which robustness analysis proceeds in economics largely conforms to Weisberg's account, as our illustration in Section 5 will make clear. However, since our claim about economics as robustness analysis concerns the robustness of results with respect to modelling assumptions, we do not require that stability analysis be performed in order for this modelling practise to be characterised as robustness analysis.[4]

By a model result we mean any proposition derivable from a model that is thought to be epistemically or cognitively important in the appropriate scientific community. In economics, typical model results are properties of equilibria, existence conditions for equilibria, and dependencies between variables derived through comparative statics. By a common structure, we mean a formal representation of the (causal) mechanism[5] which is thought to produce

---

[3] Let $\mathrm{DET}_n$ be random variables expressing whether a method of determination $n$ produces the result of interest $R$, and let RES be a random variable expressing whether the result $R$ actually holds or not. Independence of forms of determination can be defined using probabilistic independence as $\mathrm{DET}_1 \perp \mathrm{DET}_2 \ldots \perp \mathrm{DET}_n \mid \mathrm{RES}$. Note that independence required by robustness is *not* to be equated with (Bayesian) confirmational independence, which requires that the confirmation received by a hypotheses $H$ from a piece of evidence $E_1$ is independent of another piece of evidence $E_2$ (cf. Fitelson [2001]).

[4] See also (Weisberg and Reisman [2008]).

[5] As Harold Kincaid rightly pointed out to us, not all economic models include explicit descriptions of causal mechanisms. Many applied econometric models are not committed to any specific underlying mechanisms and others are purely predictive. We take equilibrium models, which constitute a great portion of theoretical modelling in economics, to represent causal mechanisms that constitutively explain system-level properties (Kuorikoski [2007]).

or constitute the phenomenon represented by the result *R*. Economists often proceed on the basis of a preliminary hypothesis or intuitive hunch that there is some core causal mechanism that ought to be modelled realistically. Turning such intuitions into a tractable model requires making various unrealistic assumptions concerning other issues. The inevitability of these idealisations and abstractions makes at least some of the assumptions of economic models always unrealistic: even a perfect economic model is idealised (Lehtinen and Kuorikoski [2007a]). In physics, it is possible, in principle, to use fundamental theories to determine how much distortion is introduced with each idealisation (cf. Odenbaugh [2005]; Weisberg [2006a]). By way of contrast, in economics there is no fundamental theory that tells the modeller which assumptions give cause for alarm and which do not and how one should go about making the models more realistic (Hausman [1985]). Real economic systems are always constituted by heterogeneous agents and characterised by changing parameter values, which implies that there are no universal or timeless constants to approximate. For most economic phenomena of interest there might not be a single true functional form, fixed over time, against which the exact form of the assumptions could be compared.

Non-economists are often annoyed by economists' seemingly sanguine attitude towards criticisms of unrealistic assumptions: such criticisms are taken seriously, i.e. published in economics journals, only if they are incorporated in an alternate formal model that shows whether, and if so how, a modified assumption changes the conclusions of the original model. The very existence of this methodological practise is evidence of the importance of uncertainty concerning the consequences of unrealistic assumptions and the concomitant importance of robustness considerations in economics (see also Gibbard and Varian [1978]). If mere unrealisticness were sufficient to invalidate a model, it would be perfectly justifiable to accept criticisms of assumptions without an accompanying formal model. Moreover, if it were easy to know which assumptions were realistic and which mattered for the model results, there would be no need to offer a formal proof that results are not robust.[6]

Because economists cannot rely on theoretical frameworks for determining the importance of various idealising assumptions, they often resort to intuitive notions of realisticness. Economic models can be made more realistic in a variety of ways. These include, but are not restricted to, taking into account a factor that was previously neglected, providing links between variables that were already incorporated into the model, restricting the domain of application, specifying in more detail institutional or other contextual factors, and

---

[6] As pointed out by Lehtinen and Kuorikoski ([2007b]), showing that some particular results are not robust with respect to a crucial assumption such as rationality immediately provides entry into even the most prestigious economics journals.

providing a more realistic account of individual behaviour by allowing deviations from rationality or incomplete information.

Turning intuitions regarding the existence of a core causal mechanism into a tractable model requires making unrealistic assumptions regarding other issues. For our purposes, we distinguish three kinds of assumptions according to the role they serve in the model: *substantial assumptions*, *Galilean assumptions*, and *tractability assumptions.* Substantial assumptions identify a set of causal factors that in interaction make up the causal mechanism about which the modeller endeavours to make important claims. Substantial assumptions are hoped to be realistic in at least two senses: the central mechanism should be at work in reality, and the 'strength' of the identified mechanism should not be minor.

For constructing models, one also needs what Cartwright ([2006]) calls Galilean idealisations (henceforth Galilean assumptions).[7] These are assumptions that serve to isolate the working of the core causal mechanism by idealising away the influence of the confounding factors (see also Mäki [1992], [1994a]).

In contrast to substantial assumptions, Galilean assumptions are unrealistic in the sense that they are thought to be false in any actual systems to which the model is applied, but they, too, are intended to have a causal interpretation: they state that a factor known or presumed to have an effect is absent in the model.

However, Galilean assumptions are typically not sufficient for deriving results from models and additional assumptions are needed to make the derivation feasible. Some modelling assumptions are thus introduced only for reasons of mathematical tractability (see Hindriks [2006]).[8] These tractability assumptions are typically unrealistic, but the falsehood they embody is hoped to be irrelevant to the model's result. Tractability requirements sometimes demand that substantial assumptions are also incorporated in ways that are more specific than desired: the causal factors making up the core mechanism have to be implemented in the model in some mathematical form, and the way in which substantial assumptions are implemented in the model may

---

[7] Ernan McMullin ([1985]) uses the term 'Galilean idealizations' more broadly to refer to various techniques of deliberate simplification in both theoretical and experimental practise. In particular it refers both to 'causal idealizations', which serve to isolate away the influence of confounding factors, as well as to 'construct idealizations', which refers to the conceptual stylization of the object or system of interest and thus in many cases also involves what we call tractability assumptions. To avoid adding yet another label to the many alternatives already available in the philosophical literature on economic models, we prefer to follow Cartwright's ([2006]) use which explicitly discusses idealisations in economic models. We also recognise that, as one of the referees pointed out (and McMullin himself admits), it may suggest a misleading picture of Galileo's scientific practise.

[8] The literature on modelling includes various closely related concepts, such as Musgrave's ([1981]) heuristic assumptions, Mäki's ([2000]) early-step assumptions, and Alexandrova's ([2006]) derivation facilitators.

introduce an element of falsehood, which is hoped to have little consequence for the result (Section 6 discusses assumptions about transport costs as an example of this). Thus a single explicitly stated modelling assumption may simultaneously encode a tractability assumption as well as a substantial assumption. Unlike Galilean idealisations, for many tractability assumptions it is often unclear what it would mean to replace them with more realistic ones: if it were possible to do without this kind of assumptions they would not be introduced in the first place (see also Weisberg [2006a]). This is why tractability assumptions are often replaced with assumptions that are also unrealistic, but in a different way.[9]

Although robustness analysis involves all three kinds of assumptions, it is only the failure of robustness with respect to tractability assumptions that is epistemically problematic, because it suggests that the result is an artefact of the specific set of tractability assumptions, which in many cases have no empirical merit on their own. If a result turns out to be robust across models that deploy different sets of tractability assumptions, but share the same set of substantial assumptions, the dependency between the latter and the result is less likely to be an artefact of the tractability assumptions. This is the function of guarding against errors of robustness analysis: it means guarding against the unknown consequences of unrealistic assumptions, the falsity of which, it is hoped, is innocuous.

Nancy Cartwright ([2006]) complains that although tractability assumptions are those that mostly need to be subjected to robustness analysis, they are also the ones for which it is rarely performed.[10] Whether she was right in claiming that this is not sufficiently done in economics is a question we cannot fully address here. Our illustration provides an instance in which tractability assumptions are also modified, and this is not an isolated example. Except for a few clear-cut cases, in practise it is hard to know in advance which assumptions are in fact tractability assumptions and which are tied to the core causal mechanism and which idealisations can be replaced with more realistic assumptions and which cannot. Robustness analysis, in its function of assessing the importance of various assumptions to the conclusions, also contributes to distinguishing between assumptions in terms of their type and role.

---

[9] Modelling assumptions can also be replaced by more general ones. For instance, consider a model in which it is assumed that variable $Y$ depends on $X$ positively and linearly; $Y=aX$, $a>0$. If a robust result is proved under the assumption that $Y$ depends positively on $X$ ($Y=f(X), f'>0$), the new model tells us that the (often suspicious) linearity property was in fact not needed for the result. Here robustness analysis both guards against error and facilitates assessment of the importance of different model components: it shows that linearity did not lead us astray, and that what was needed to get the result was only the assumption that $Y$ increased with $X$.

[10] Cartwright does not use the term 'tractability assumptions'. She calls assumptions needed for the derivation to go through 'non-Galilean idealisations'.

The way in which we use the term ‘robustness analysis’ does not always coincide with the way in which economists use it: they sometimes use the term to refer to particular methods such as perturbation and sensitivity analysis, and they do not always talk of robustness analysis when they present alternate models in which one or more tractability assumptions are modified. For our purposes, for the comparison between alternative models to qualify as derivational robustness analysis, it is sufficient (and necessary) that the different models share a common structure and deploy different tractability assumptions, so that inferences regarding the dependency of the results on the various model components can be made. For an activity to count as robustness analysis, it thus need not be intentionally conducted by an individual economist. The modified models are often, although not necessarily, presented by different economists than the one(s) who proposed the original model (as shown below, this holds for our case study). In this sense, then, our claim is that theoretical model building in economics is to be understood as *collective derivational robustness analysis*.

## 4 The Epistemic Import of Robustness Analysis

The common view about derivational robustness is that it is empirically vacuous and that it constitutes a suspicious form of pseudo confirmation (Cartwright [1991]; Orzack and Sober [1993]). Obviously empirical data are the natural arbiters for evaluating the validity of models and theories. Even so, in the absence of access to the kind of data that would straightforwardly have a bearing on whether a certain modelling result was accurate, it may be justifiable to use robustness analysis. Since economic models are based on unrealistic assumptions and some branches of economics have been criticised for a lack of empirical testing (Blaug [1980]; Green and Shapiro [1994]), the accusation of neglect of empirical evidence in favour of empirically uninformed model manipulation is indeed more relevant in economics than in many other fields. In this section, we defend the practise of robustness analysis against the accusation of epistemic sterility.

According to Robert Sugden ([2000]), since robustness analysis is a matter of comparison between models, it does not license ‘the inductive leap’ from models to economic reality. Similarly, Orzack and Sober ([1993], p. 539) maintain that ‘it is worth considering the possibility that robustness simply reflects something common among the frameworks and not something about the world those frameworks seek to describe’. The fundamental worry is that derivational robustness analysis only teaches us about the properties of models and therefore that we are not justified to draw any conclusions about things outside of them.

Orzack and Sober provide the following characterisation of possible epistemic states concerning a set of models with respect to which robustness analysis is to be conducted:

(i) We know that one of a set of models $M_1, \ldots ,M_n$ is true.

(ii) We know that each of the models $M_1, \ldots ,M_n$ is false.

(iii) We do not know whether any of models $M_1, \ldots ,M_n$ is true.

They acknowledge that an instance of case (i) would prove that the common result $R$ is true,[11] but claim that this is not the usual case. In cases (ii) and (iii) it is unclear why the fact that $R$ is a joint prediction (i.e. a robust result) of the models should be regarded as evidence for $R$'s truth, given that robustness may merely reflect something the models share, thus having no implications about the world outside the models. As a logical point about deduction their argument is valid: the robustness of a result does not in itself guarantee that it is true.

The way in which Orzack and Sober present their criticism suggests that they take models as a whole, rather than their components or aspects (cf. Mäki [2006]) as either true or false. If falsity means the absence of the-whole-truth as well as of nothing-but-the-truth,[12] then models are always false because they always contain idealisations. Hence, only case (ii) seems to be relevant. Orzack and Sober provide the following illustration of case (ii): all biological models, in which natural selection is the only force acting on a population, have the consequence that the population size is infinite.[13] They refer to the infinite population size as a 'result'. If a result is allowed to be any proposition that can be derived from the model, Orzack and Sober's example boils down to the idea that falsities due to idealisations may always engender falsities in results. But awareness of this problem is precisely the reason to engage in robustness analysis. The results relevant for robustness analysis are those having to do with the substantial assumptions. Sober and Orzack's illustration provides a poor counterexample against the epistemic importance of robustness analysis because it does not deal with a robust derivation of an empirically interpretable result that could be caused by the mechanism defined by the substantial assumptions. That natural selection is the only force acting on a population could not conceivably be the cause for any population being infinite and infiniteness of the population is not the result of interest in these models. Instead, the example correctly points out that, for

---

[11] And Cartwright ([1991]) would have added that the robustness of the result would still not tell us which is the true model.

[12] This distinction was proposed in (Sen [1980]). See (Mäki [1992], [1994a]) for further analysis.

[13] It thus seems that here they are speaking about the truth of an individual assumption rather than that of a model.

mathematical reasons, models with a certain substantial assumption also have a certain tractability assumption (namely infinite population).

We argue instead that the relevant case, which Orzack and Sober do not contemplate, is one in which we know that all the models are false (in that they contain false assumptions), but some of their elements may nonetheless be true. A single model may have both true and false elements (Mäki [2004], [2009], [forthcoming]; Hindriks [2008]). When we refer to a model's elements, we mean individual assumptions, results, and the model's theoretical claim that relates the core mechanism to the results of the model. The truth values of these elements need not be the same. For example, many assumptions could be false and the theoretical claim could still be true.

Before conducting robustness analysis we do not know for sure which part of the models is responsible for the result, although modellers usually have strong intuitions about this issue. If a result is implied by multiple models, each containing different sets of tractability assumptions, we may be more confident that the result depends not on the falsities we have introduced into the modelling, but rather on the common components (Weisberg [2006b]). Robustness analysis thus increases our confidence in the claim that the modelling result follows from the substantial assumptions, i.e. that some phenomenon can be caused by the core mechanism. This is what Levins and Weisberg call the robust theorem.

What we learn from robustness analysis indeed concerns only the properties of the models. The point is that if our degree of belief in the truth of different modelling assumptions varies, then learning about the model properties may justifiably change our degree of belief in the results. If we are initially more confident about the substantial assumptions than we are about the tractability assumptions, then learning that our modelling results do not crucially depend on the tractability assumptions increases our degree of belief in the result concerning the effects of the modelled mechanism. There is thus nothing dubiously non-empirical or Münchausen-like in the epistemic import of derivational robustness.

Making sense of this epistemic import thus requires looking at and attributing different degrees of credibility or reliability to *parts* of models. In order to illustrate this more clearly, let us look at another argument by Orzack and Sober against the epistemic importance of derivational robustness, this time concerning what happens when analytic models encounter data. Here is how they formulate it:

> Suppose that each of two competing models is reasonably well supported by the data. If R is a robust theorem that they share, should we conclude that the data support the common element in the two models? Presumably, if the data had been different, we would not have regarded the

> models as well supported. The question is whether we would be prepared to doubt R in this circumstance as well. If not, then this robust theorem is not tested by the data and consequently is not well supported by them […] the robustness of R is not by itself a reason to believe it. Whether R is plausible depends on the data and not on the fact that R is robust […] Testability of predictions […] depends upon having nonrobust theorems to test, that is, those that are not entailed by all of the models under test. (Orzack and Sober [1993], p. 541)

Their claim is that it is the data rather than robustness that do the confirming, or else the result is simply untestable and thus empirically empty. As stated above, robustness analysis does not in itself provide empirical confirmation. However, their argument rests on an ambiguity: they do not specify what is meant by the claim that two models are supported by some data. Orzack and Sober seem to take the Duhem–Quine thesis literally, assuming that a theory or a model is confirmed or disconfirmed only as a whole—the only question being whether it fits the data. To claim that a model reveals anything interesting about the phenomenon on the basis of its fit to a set of data, however, would require an additional inferential step.

If two models are in competition but cannot be empirically distinguished, i.e. if they stand and fall together given the kind of data available, then the only thing that *can* be tested against the data is what they have in common, including the robust theorem. Robustness analysis is thus necessary for determining which *part* of the model is tested or confirmed (Wimsatt [2007], p. 53). Orzack and Sober's example thus shows how robustness considerations are crucial in determining what can be confirmed with given data. Confirmation may concern the assumptions of the model, or the consequences of those assumptions. We are often interested in knowing whether the causally important parts of our models, rather than the unimportant tractability assumptions, are confirmed or disconfirmed. If there were a direct empirical test of whether the result of interest holds, conducting it would obviously be the thing to do. Robustness analysis is about the assessment of the security of our inferences from assumptions of which we do not know whether they are true and is relevant precisely when there is no direct way of empirically ascertaining whether the conclusions of our inferences are true.

Although robustness analysis does not provide a means of *finding* causally important mechanisms, it serves to distinguish them from irrelevancies (see also Odenbaugh [2005]). When a scientist constructs a model, he or she tries to incorporate the important causal factors and leave out the irrelevancies. If the model fails to depict the important factors, it will not give a true view of the situation even if its results are robust with respect to various tractability assumptions. The contribution that robustness analysis provides is that it allows the modeller to be just a little bit more certain that what he or she hoped was irrelevant is, in fact, irrelevant. As Levins puts it, it serves to determine the extent to which we can get away with *not* knowing all the details, but still understand the system (Levins [1993], p. 554).

Robustness failure in a modelling result with respect to an assumption could, in principle, always be interpreted in two alternative ways: the lack of robustness either demonstrates a shortcoming in the modelling framework (a tractability assumption drives the result) or it suggests a new empirical hypothesis about a causally relevant feature in the modelled system. In the course of model refinement, what was thought to be a tractability assumption may be re-interpreted as a Galilean or even a substantial assumption and *vice versa*. The decision on how to interpret the failure of robustness with respect to an assumption is often not that difficult to make. This decision may be based on the prior confidence assigned to the problematic assumption and on whether the dependency of the result on this assumption can even be given a sensible empirical interpretation. Nevertheless, the open-endedness and fallibility of this decision has to be acknowledged.

## 5 An Illustration: Geographical Economics Models

Geographical economics (henceforth GeoEcon) is a recent approach to spatial issues developed within economics, the aim of which is to explain the spatial location of economic activity. Paul Krugman ([1991]) provided the first model, generally referred to as the CP model (C for 'core' and P for 'periphery').[14] Following its appearance, a growing body of theoretical literature has refined and extended Krugman's original model. The CP model as well as its followers (summarised in Fujita et al. [1999]) depends crucially on a set of unrealistic assumptions, or 'modelling tricks', as geographical economists call them. In order to turn a set of interesting examples, or a collection of special cases, into a general theory of the location of economic activity, geographical economists engage in what we have identified as robustness analysis: many of the subsequent GeoEcon models appear to be checking whether the main conclusions of the CP model remain valid when some of its unrealistic assumptions are altered. Let us begin by briefly setting out the main ingredients of the CP model, and we will then look at a few GeoEcon models that explore the robustness of its results.

Krugman ([1991]) employs the Dixit and Stiglitz ([1977]) general equilibrium model of monopolistic competition with transportation costs and labour mobility to derive a core-periphery pattern, in other words a situation in which the bulk of economic activity is located in one region. The model operates under the following assumptions: there are two regions, identical in all respects, and

---

[14] In 2008 Paul Krugman was awarded the Nobel Prize in economics for his seminal contribution to the field of geographical economics.

two sectors in the economy. The perfectly competitive (agricultural) sector employs unskilled labour which is distributed equally between the two regions and cannot move across them. The monopolistically competitive[15] (manufacturing) sector uses only one input, skilled labour, which can move across regions, and each firm produces a variety of a differentiated product (one per firm). Consumers love variety, that is, their utility increases not only with the amount of a given variety consumed but also with the number of varieties available at their location. This preference is expressed by a constant elasticity of substitution (CES) utility function that is symmetric in a bundle of differentiated products.[16] The trade of the manufactured good produced is subject to transportation costs. In order to avoid modelling a separate transportation sector, the cost of transporting goods is assumed to be of the Samuelsonian iceberg form: a fraction of the good transported melts away in transit.

In this setting, the distribution of the manufacturing sector across the two locations is determined by the balance between centripetal and centrifugal forces. The centripetal forces arise from a positive feedback process: the more firms and workers there are in the region, the more attractive the region becomes for further firms and workers (market-size effect). Due to economies of scale and transportation costs, firms have an incentive to locate in the larger market for their product so as to economise on transportation costs, and workers/consumers tend to locate where the producers of manufactured goods are in order to economise on living costs and to benefit from a larger variety of goods (recall that each firm produces a different variety). The centrifugal forces arise from the need to serve the immobile factor, which remains equally distributed across the two regions, and by competition effects: in the larger region, firms face tougher competition and higher input prices. Figure 1 provides a stylized depiction of the market-size and competition effects.

In succinct form, the GeoEcon core causal mechanism can be formulated as follows (see Mäki and Marchionni [2009]):

> [$C_{\mathrm{CP}}$] In the presence of immobile and mobile activities, the interaction among economies of scale, monopolistic competition and transportation costs gives rise to centripetal and centrifugal forces.

---

[15] Manufacturing firms are assumed to benefit from economies of scale: production costs per unit of the product decrease as the size of production increases. If there are no economies of scale, firms would have an incentive to split production across locations, making firms' locational choices trivial (that is, without economies of scale firms need not choose whether to locate in one or the other region, for they can locate in both). The assumption of economies of scale however is incompatible with the standard economic assumption of perfect competition. The monopolistic competition model provides a means to model economies of scale within a framework of general equilibrium.

[16] CES utility functions are a commonly used class of utility functions of which the Cobb–Douglas is a special case. Another common utility function is quadratic, used in later modelling efforts to check how much the CP results depend on expressing utility in terms of CES functions.

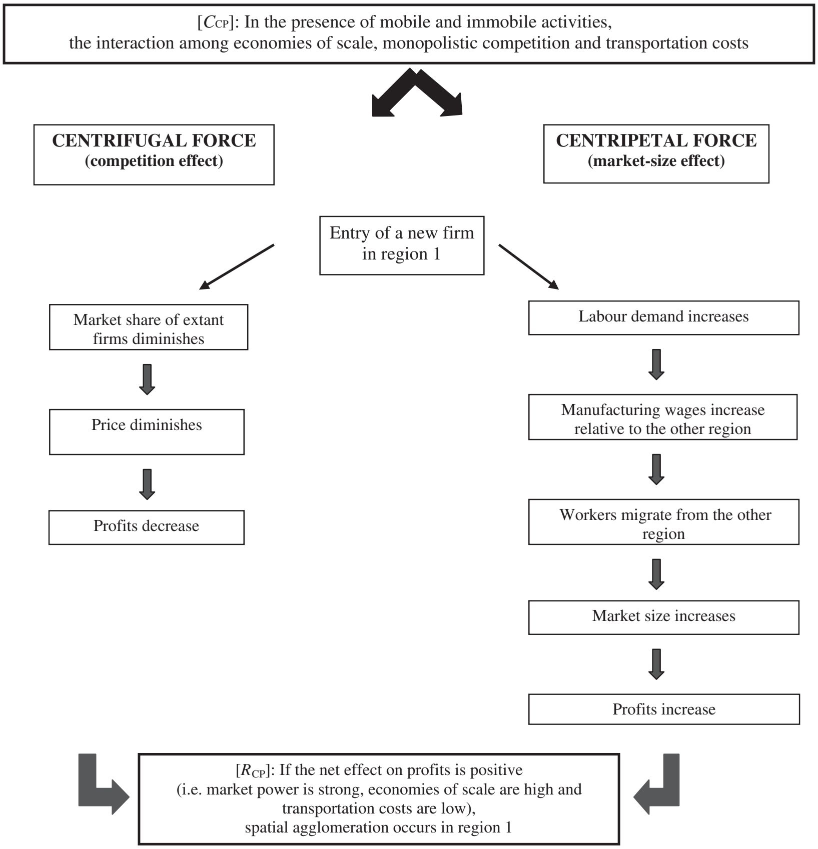

**Figure 1.** A stylised description of the competition and market-size effects in the CP model of GeoEcon. The entry of a new firm in a region has two opposite effects. If the net effect on profits is positive, all the mobile activity will be clustered in one region (modified from Capello [2004], p. 306).

The result of the model is as follows:

> [$R_{CP}$] Ceteris paribus, spatial agglomeration occurs when economies of scale are high, market power is strong, and transportation costs are low (that is, when the centripetal forces are stronger than the centrifugal forces).

Lower transportation costs, strong market power, and high economies of scale reduce the effects of market crowding vis-à-vis the market-size effect, and hence centripetal forces are stronger than the centrifugal forces.[17] Signif-

---

[17] Lower transportation costs make serving the immobile workers in the other region less expensive, strong market power reduces the effects that the entry of new firms has on prices, and the higher the economies of scale, the more advantageous it is to be located in the larger market.

icantly, the CP model shows that changes in transportation costs affect the balance between centripetal and centrifugal forces in a non-linear way (see Ottaviano [2003]).[18]

The CP model relies on very specific functional forms of utility (namely, CES functions), production (a homothetic function with skilled labour as the only input), and transportation technology (the Samuelsonian iceberg form). These assumptions are made mainly for reasons of tractability, and it is hoped that the results [$R_{\text{CP}}$] do not crucially depend on them.[19] Later models explore the robustness of the results of the CP model. In particular, Ottaviano et al. ([2002]) employ quadratic utilities instead of CES functions and assume linear instead of iceberg transportation costs. In contrast to the CP model, these assumptions entail that demand and substitution elasticities vary with prices and that equilibrium prices depend on the distribution of firms and consumers. The main conclusions of the CP model are found to be robust with respect to these changes. Ottaviano et al. ([2002], p. 432) thus conclude:

> The main results in the literature do not depend on the specific modelling choices made, as often argued by their critics. In particular, the robustness of the results obtained in the CP model against an alternative formulation of preferences and transportation seem to point to the existence of a whole class of models for which similar results would hold.

A series of models (Forslid and Ottaviano [2003]; Ottaviano and Robert-Nicoud [2006]; Ottaviano [2007]) labelled Footloose Entrepreneurs (FE) derive a core-periphery pattern driven by the mobility of labour, as in the CP model, but assume a different specification of the production function. Whereas the CP model uses a homothetic function in which only skilled labour appears, FE models use the Flam and Helpman ([1987]) functional form which assumes that firms use both skilled and unskilled labour. This modification has implications for other functional relationships in the model: in the CP model the number of firms functions as the adjustment variable, whereas in the FE models the number of firms varies with wage. Nonetheless, the FE models derive the same qualitative result as the CP models: [$R_{\text{CP}}$].

---

[18] The CP model has two other important qualitative features: (i) catastrophic agglomeration: a small change in the critical parameters can tip the economy from a situation of dispersion to one of full agglomeration, and (ii) locational hysteresis or path-dependency: transitory shocks can have permanent effects. These features are generally regarded as distinctive of GeoEcon. Puga ([1999]), however, shows that in the presence of congestion in the agglomerating region or of heterogeneity across firms, the transition from dispersion to agglomeration becomes gradual. Ottaviano et al. ([2002]) predict catastrophic agglomeration without hysteresis. Although we do not discuss these models here, their contributions vis-à-vis the CP model are also to be interpreted as robustness analysis.

[19] Even with these simplifications, deriving the results of the CP model requires numerical computations.

Despite modifications in some components of the CP model, one set of ingredients remains invariant across alternative frameworks: the presence of economies of scale and imperfect competition, of transportation costs, and of mobile and immobile activities. This set of ingredients together with their interaction corresponds to what we call substantial assumptions and defines the core causal mechanism [$C_{\mathrm{CP}}$]. That is, the interaction between these ingredients generates the centrifugal forces (the market-crowding effect due to increased competition) and the centripetal forces (the market-size effect due to economies of scale and transportation costs), the relative strength of which determines whether or not agglomeration occurs. Since these substantial assumptions are maintained in models carrying different tractability assumptions, they (rather than the tractability assumptions) are believed to be responsible for the derivation of the same result [$R_{\mathrm{CP}}$].

Note that, as discussed in Section 3, substantial assumptions need to be introduced into the model in some specific mathematical form, which may embody a degree of falsehood. Take transportation costs, for example. In the CP model, they feature as iceberg transportation costs, an assumption that geographical economists admit is highly unrealistic but convenient, for it allows them to avoid modelling a separate transportation sector. Transportation costs nevertheless represent a crucial component of the core causal mechanism and are indeed incorporated in some way or another in all of the models. That the same results obtain with alternative specifications of transportation costs suggests that the results crucially hinge not on the unrealistic assumption of iceberg transportation costs but on the realistic substantial assumption that goods are costly to transport.

The different models therefore share a robust result [$R_{\mathrm{CP}}$] and a set of substantial assumptions [$C_{\mathrm{CP}}$].[20] This gives us the theoretical claim of this family of GeoEcon models: *Ceteris paribus*, if firms benefit from economies of scale, goods are costly to transport, and there are both immobile and mobile activities, spatial agglomeration occurs when economies of scale are high, market power is strong, and transportation costs are low. This is what Levins and Weisberg call a 'robust theorem'.

Robustness analysis also provides information when the result turns out not to be robust. Ottaviano and Thisse ([2004]) show how more realistic models, including an additional spatial cost (such as congestion or transport costs for goods in the agricultural sector), demonstrate that the value of the transport cost relative to this additional spatial cost is crucial in determining whether dispersion or agglomeration occurs. For instance, adding the transport costs

---

[20] Not all model assumptions have been checked for robustness, which makes the exclusive dependency of the result on the substantial assumptions less certain. As explained below, the epistemic import of robustness analysis is a matter of degree.

of the agricultural good gives a different result from [$R_{\mathrm{CP}}$]. In this model, when this cost is low, agglomeration occurs at intermediate levels of the transport cost of the manufactured good, exactly as predicted in the CP model. However, when the cost of transporting the agricultural good is high, the industry is always dispersed at high levels of the cost of transporting the manufactured good.

In this case the divergence in results does not invalidate the theoretical claim that connects [$C_{\mathrm{CP}}$] to [$R_{\mathrm{CP}}$] in the CP model. First, dispersion (and not agglomeration) now occurs when the transportation costs of the manufactured good are low because the price differential of the agricultural good constitutes a further centrifugal force that was absent in the CP model. More importantly, when transportation costs are high, the centrifugal force (the crowding effect on the market) is exactly the same as in the CP model. This confirms the fact that, in the absence of a significant additional cost, the relationship between high transport costs of the manufactured good and dispersion is as predicted in the CP model. In the absence of an additional spatial cost, the connection between [$C_{\mathrm{CP}}$] and [$R_{\mathrm{CP}}$] is robust. Second, the finding that industry is always dispersed when the costs of shipping the traditional good is high is taken to show that 'the level of the agricultural good's transport costs matters for the location of industrial firms' (Ottaviano and Thisse [2004], p. 35).The breakdown of [$R_{\mathrm{CP}}$] is interpreted as implying that the factor in question, transportation costs for the agricultural good, which was previously neutralised by means of a Galilean assumption, affects the results and constitutes a further centrifugal force that was absent in the CP model. This alternative is used to study how making a more realistic model by relaxing this particular Galilean assumption affects the conclusions about the working of the mechanism.

It is instructive to contrast this case, in which a Galilean assumption is made to neutralise the effect of a factor presumed to have an impact, with the previous ones, which focused on the replacement of tractability assumptions. Whereas in this case the relaxation of the Galilean idealisation makes for a more realistic model, in the previous ones it is unclear whether for instance the tractability assumption of quadratic utility functions is more realistic than the assumption of CES functions—they are just different. The connection between [$C_{\mathrm{CP}}$] and [$R_{\mathrm{CP}}$] is thus shown to be robust across models with different tractability assumptions that cannot be ranked according to their realisticness. The two cases allow us to appreciate the two intertwined functions of robustness analysis: it protects against unknown errors of particular falsehoods (this is especially evident in the case of replacing different kinds of falsehoods), and it tracks the relative importance of various components of theoretical models (this is especially evident in the case of replacing assumptions with more realistic ones).

## 6 Independence of Derivations

The discussion has thus far centred on the role of robustness analysis in assessing the relative importance of modelling assumptions. What needs to be established is that derivational robustness analysis is a species of general robustness analysis in the sense discussed by Wimsatt and that the same epistemic rationale applies to it.

Robustness confers a higher degree of belief on a result if the methods of determination with respect to which the result is robust are independent of each other. Thus, as Orzack and Sober ([1993]) point out, the intelligibility of the concept of robustness crucially depends on the possibility of giving an account of the independence of the methods of determination. Recall that here we are concerned with derivational robustness, that is, the robustness of a theoretical result with respect to different modelling assumptions. Orzack and Sober seem to take it for granted that independence in derivational robustness is a matter of independence *between models*—but what does independence between models mean? Orzack and Sober address two candidates: what they call *logical* and *statistical* independence. Two models are logically independent if neither implies the truth or falsity of the other, and they are statistically independent if we can select one model from a sample so that the probability of selecting that model is independent of the probability of selecting another. They quickly abandon both options and for good reasons: models are very seldom logically independent of each other, and assigning probabilities to the truth of them does not seem reasonable.

In our opinion, Orzack and Sober have not successfully shown that it is impossible to provide an account of independence that is relevant to robustness analysis. It is important to realise that even though the various models in a set $M$ are not independent because they share some assumptions, it is the *independence of individual tractability assumptions within a set of similar models* that is crucial for derivational robustness, rather than the independence of models. It is usually not the independence of different tractability assumptions with respect to each other *within a single model* that matters, but rather the independence of assumptions that have a similar role in a *set* of models. As Wimsatt ([1980a], p. 310) notes, the models must be similar so that we can compare them and isolate their similarities and differences.

Orzack and Sober are right in claiming that Levins' notion of robustness does not really incorporate the independence assumption. Levins ([1966], [1968]) notes that the various models share the biological assumption, but beyond that, he does not say much about what the common part of a set of models is supposed to be. It may consist of a specification of a mechanism, or perhaps even a fully fledged theory. Something similar is going on in economic models. As exemplified in the GeoEcon case, different models share some but not all substantial economic assumptions, possibly those that specify the core mechanism.

Let us express Levins' conception of robustness as follows; let $C\&V_i$ denote a model $M_i$ that is based on combining the common part shared by a set of models $M$ with the specific part $V_i$. Let $A \vdash \mathrm{B}$ denote '$B$ is derivable from $A$' (within some standard formal system). Robustness of the relationship $C \vdash R_{\mathrm{M}}$ requires that

$(C\&V_1) \vdash R_{\mathrm{M}}$, and

$(C\&V_2) \vdash R_{\mathrm{M}}$, and ...

$(C\&V_n) \vdash R_{\mathrm{M}}$.

Robustness of a derivation is hardly interesting if $R_{\mathrm{M}}$ can be derived without making any tractability assumptions. It is therefore natural to think that $\sim(C \vdash R_{\mathrm{M}})$, although this is not a formal requirement. Levins' treatment of robustness gives the impression that the set of tractability assumptions must be exhaustive, but we think that this is too restrictive.[21] In practise the typical case is one in which it is not possible to define an exhaustive set of possible tractability assumptions, let alone go through the whole possibility space. Indeed, the difficulty of modifying even somewhat easily specifiable tractability assumptions is one reason for thinking that economic model building qua robustness analysis is not a trivial activity. Allowing for a non-exhaustive set implies that we cannot be sure that the relationship is actually robust, but it does not remove the epistemic relevance of robustness analysis altogether.[22] Its epistemic import comes in degrees.

A robust theorem works on the assumption that the result $R_{\mathrm{M}}$ depends on some central mechanism $C$. Levins' ([1966]) unclear but intuitively appealing claim that 'our truth is the intersection of independent lies' could be taken to mean that result $R_{\mathrm{M}}$ can be derived from mechanism-description $C$ using multiple independent sets of untrue tractability assumptions. Various falsities involved in the different derivations do not matter if robustness analysis shows that result $R_{\mathrm{M}}$ does not depend on them. In the GeoEcon

---

[21] Levins ([1993]) proposes evaluating robustness according to the following framework. Let $R_{\mathrm{M}}$ denote a robust theorem. Let $C$ denote the common part of all models in some set of models $M$. Let $V_i$ denote the variable part of the models. Then, if $CV_1+CV_2+\ldots=C(V_1+V_2+\ldots)\rightarrow R_{\mathrm{M}}$, and if the set of $V_i$s is exhaustive, then $(V_1+V_2+\ldots)=1$, and $C\rightarrow R_{\mathrm{M}}$. Since Levins does not define the addition sign (+) in this formalism, it is difficult to tell what exactly is being claimed.

[22] As Weisberg shows, the Volterra principle was eventually mathematically proven to be independent of problematic auxiliary assumptions (Weisberg [2006b]). We believe this kind of case to be the exception rather than the norm.

case, changing the specifications of the functional forms and technologies while retaining the assumptions depicting the core causal mechanism, [$C_{\mathrm{CP}}$], yielded the same qualitative conclusions [$R_{\mathrm{CP}}$]. The derivation of [$R_{\mathrm{CP}}$] from different sets of unrealistic assumptions increases confidence that [$R_{\mathrm{CP}}$] depends on [$C_{\mathrm{CP}}$] rather than on the GeoEcon modelling tricks.

What kind of independence of assumptions, then, is relevant for derivational robustness? Suppose one assumption is that transportation costs are iceberg-like and another is that they are linear. The two are clearly not logically independent since they both cannot be true of the same system. Again, Wimsatt's account is helpful here: what is needed is an account of how the different tractability modelling assumptions could be thought of as not sharing the same biases and other sources of possible error. As stated in Section 2, robustness in general requires that the success of methods of determination be independent, given the knowledge of whether the result holds.[23] If we generalise this idea to the world of models we find that whether the result can be derived with a given set of tractability assumptions should be independent of whether it can be derived with an alternative one. The problem, of course, is that whether the result can be derived from the substantial assumptions about the core causal mechanism and a particular set of tractability assumptions is not a random variable. Whether a relation of derivability holds cannot be a matter of probabilities. Talk of reliability, biases, and errors seems out of place here because we are dealing with logical relations of formal derivability, not causal processes of measurement or experimentation.

Our proposal for solving this problem is to go subjectivist and relativise the epistemic gain from robustness to the epistemic situation of the modeller or of the relevant scientific community at a certain time. Modelling can be considered as an act of inference from a set of substantial assumptions to a conclusion (see also Kuorikoski and Lehtinen [2009]). Tractability assumptions are typically needed for the process of inference to be feasible, but these assumptions may induce errors in the modelling process: they may lead us to believe falsities about the world even if the substantial assumptions are true. By errors we mean the unwanted consequences of convenient but literally false assumptions rather than logical or mathematical blunders in the derivation. Our notion of independence ought to satisfy the requirement that robustness analysis decreases the probability of making such false inferences. We thus propose that the modeller should have no positive reason to believe that if one tractability assumption induces a certain kind of error (due to its

---

[23] Robustness in general requires that ($\mathrm{DET}_1 \perp \mathrm{DET}_2 \ldots \perp \mathrm{DET}_n | \mathrm{RES}$) (see footnote 3). If we plug in the models, we get ($(C\&V_1) \vdash R_{\mathrm{M}} \perp (C\&V_2) \vdash R_{\mathrm{M}} \ldots \perp (C\&V_n) \vdash R_{\mathrm{M}} | R_{\mathrm{M}}$). The problem is that the ($(C\&V_i) \vdash R_{\mathrm{M}}$) are not random variables.

falsehood) in the result, so does another one.[24] Given that the modelling result of interest ($R_M$) is correct, prior probabilities concerning whether $R_M$ can be derived from $C\&V_1$ or $C\&V_2 \dots C\&V_n$ should be (roughly) independent. If the probabilities are independent in this way, then observing that the models lead to the same result rationally increases our degree of belief in the result.

In the GeoEcon case, if assuming CES and quadratic utility functions had exactly the same implications, then they would induce the same kind of falsehood in the model. However, quadratic utility functions entail that demand and substitution elasticities vary with prices and that equilibrium prices depend on the fundamentals of the market, whereas CES functions entail constant elasticity of demand and substitution and the independence of prices from the fundamentals of the market. These different assumptions are both literally false, but they create different kinds of distortions in the models.

In sum, economic models in a given sub-field are often far from independent because they differ only with respect to a few particular modelling assumptions. Robustness analysis in economics is thus usually a special, degenerate form of general robustness analysis as Wimsatt defines it: checking the robustness of a result with respect to a limited set of modelling assumptions that are usually plainly unrealistic. If a result is robust with respect to particular tractability assumptions, the empirical falsity of these assumptions does not provide grounds for criticising it. As indicated above, the value of derivational robustness analysis lies in rooting out error in our inferences, from diverse and uncertain assumptions to conclusions, and thereby increasing our confidence in the conclusions. If the substantial assumptions had no empirical merit in the first place, then robustness clearly would have no epistemic value—it would be just a similarity relation between members of a peculiar set of abstract entities. The epistemic relevance of robustness thus depends on there being realistic assumptions. It would be possible to deny its epistemic value only by denying that substantial assumptions are ever realistic in economics. Although this is a remotely plausible view that may even have been entertained, it is clearly a rather extreme position.

## 7 Concluding Remarks

The practise of economic theorising largely consists of building models under slightly different assumptions yielding familiar results. We have argued that this practise makes sense if it is seen as derivational robustness analysis.

---

[24] We do not wish to downplay the practical epistemic difficulties of establishing independence, nor the consequences of failure to detect dependence. Wimsatt argues, for example, that even though causal inertness of group selection is a common result of multiple different evolutionary models, it is not robust since all they share a common implicit reductionist bias and are thus not independent. Impossibility of group selection is, in fact, an artefact of this implicit and shared modelling assumption (Wimsatt [1980b]).

Robustness analysis is a sensible epistemic strategy in situations in which we know that our assumptions and inferences are fallible, but not how fallible and in what way. It is difficult in economics to subject models to straightforward empirical tests in that the theory does not necessarily indicate which idealisations are truly fatal or crucial for the modelling result, and the nature of the phenomena studied implies that idealisations cannot be completely eliminated. This is because economics differs from (at least some subfields of) physics in that the systems it studies are fundamentally heterogeneous, open, and continually changing (Hausman [1981]).

If things were different, i.e. if economic systems were grounded on invariant natural constants that could provide secure foundations, the best way to increase the reliability of economic theory would be to eliminate errors concerning the axioms that describe them (see also Boyd and Richerson [1987], p. 402; cf. Rosenberg [1992], Chapter 5). Since this is not the case, the best way to increase the reliability of economic theory is to control for the effects of the inevitable errors, not to pretend that they do not exist. This explains why none of the so-called fundamental axioms is truly sacrosanct and also why empirical refutation of some of the standard economic modelling assumptions need not be devastating for the whole enterprise. As Daniel Hausman has noted, economists are willing to replace any constituent of the 'basic equilibrium theory' with different alternatives (Hausman [1992a], p. 31; Hausman [1992b], p. 52). In order to control for the epistemic consequences of the inevitable errors in modelling assumptions, the theoretician is free to violate any axioms he or she pleases, as long as it is done in a way that helps to *locate* the sources of possible error that are relevant to the result of interest, i.e. with an explicit formal model showing how the result depends on particular problematic assumptions.

Derivational robustness analysis guards against errors in theorising when the problematic parts of the methods of determination, i.e. the sets of tractability assumptions, are independent of each other. In economics, proving robust theorems from different models with diverse unrealistic assumptions helps us to evaluate which results correspond to important economic phenomena and which ones are merely artefacts of particular tractability assumptions. Although derivational robustness does not provide empirical confirmation, learning that a modelling result is robust increases confidence in the reliability of the result in circumstances in which our confidence in the different assumptions varies.

## Acknowledgements

We wish to thank Kevin Hoover, Harold Kincaid, Bauke Visser, the Philosophy of Science group at University of Helsinki, audiences at MS2 and SPSP conferences and at the Erasmus Institute for Philosophy and Economics research seminar, as well as several anonymous reviewers for useful comments on the manuscript. The usual disclaimer applies.

*Department of Social and Moral Philosophy*
*P.O. Box 9, 00014 University of Helsinki, Finland*
*jaakko.kuorikoski@helsinki.fi*
*aki.lehtinen@helsinki.fi*
*caterina.marchionni@helsinki.fi*

## References

Alexandrova, A. [2006]: 'Connecting Economic Models to the Real World: Game Theory and the FCC Spectrum Auctions', *Philosophy of the Social Sciences*, **36**, pp. 173–92.

Blaug, M. [1980]: *The Methodology of Economics: or, How Economists Explain*, Cambridge: Cambridge University Press.

Bogen, J. and Woodward, J. [1988]: 'Saving the Phenomena', *The Philosophical Review*, **97**, pp. 303–52.

Bovens, L. and Hartmann, S. [2003]: *Bayesian Epistemology*, Oxford: Clarendon.

Boyd, R. and Richerson, P. [1987]: 'Simple Models of Complex Phenomena: The Case of Cultural Evolution', in J. Dupré (*ed*), *The Latest on the Best*, Cambridge, MA: MIT Press, pp. 27–52.

Capello, R. [2004]: *Economia Regionale*, Bologna: Il Mulino.

Cartwright, N. [1991]: 'Replicability, Reproducibility, and Robustness: Comments on Harry Collins' *History of Political Economy*, **23**, pp. 143–55.

Cartwright, N. [2006]: 'The Vanity of Rigour in Economics: Theoretical Models and Galilean Experiments', in P. Fontaine and R. J. Leonard (*eds*), *The 'Experiment' in the History of Economics*, London: Routledge, pp. 135–53.

Dixit, A. K. and Stiglitz, J. E. [1977]: 'Monopolistic Competition and Optimum Product Diversity', *The American Economic Review*, **67**, pp. 297–308.

Fitelson, B. [2001]: 'A Bayesian Account of Independent Evidence with Applications', *Philosophy of Science*, **68**, pp. 123–40.

Flam, H. and Helpman, E. [1987]: 'Vertical Product Differentiation and North-South Trade', *The American Economic Review*, **77**, pp. 810–22.

Forslid, R. and Ottaviano, G. I. P. [2003]: 'An Analytically Solvable Core-Periphery Model', *Journal of Economic Geography*, **3**, pp. 229–40.

Fujita, M., Krugman, P. and Venables, A. J. [1999]: *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.

Gibbard, A. F. and Varian, H. R. [1978]: 'Economic Models', *Journal of Philosophy*, **75**, pp. 664–77.

Green, D. P. and Shapiro, I. [1994]: *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*, New Haven: Yale University Press.

Hacking, I. [1983]: *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge University Press.

Hausman, D. M. [1981]: 'John Stuart Mill's Philosophy of Economics', *Philosophy of Science*, **48**, pp. 363–85.

Hausman, D. M. [1985]: 'Is Falsificationism Unpractised or Unpractisable?', *Philosophy of the Social Sciences*, **15**, pp. 313–9.

Hausman, D. M. [1992a]: *Essays on Philosophy and Economic Methodology*, Cambridge: Cambridge University Press.

Hausman, D. M. [1992b]: *The Inexact and Separate Science of Economics*, Cambridge: Cambridge University Press.

Hindriks, F. A. [2006]: 'Tractability Assumptions and the Musgrave–Mäki Typology', *Journal of Economic Methodology*, **13**, pp. 401–23.

Hindriks, F. A. [2008]: 'False Models as Explanatory Engines', *Philosophy of the Social Sciences*, **38**, pp. 334–60.

Hoover, K. D. and Perez, S. J. [2000]: 'Three Attitudes towards Data Mining', *Journal of Economic Methodology*, **7**, pp. 195–210.

Hoover, K. D. and Perez, S. J. [2004]: 'Truth and Robustness in Cross-Country Growth Regressions', *Oxford Bulletin of Economics and Statistics*, **66**, pp. 765–98.

Krugman, P. [1991]: 'Increasing Returns and Economic Geography', *The Journal of Political Economy*, **99**, pp. 483–99.

Kuorikoski, J. [2007]: 'Explaining with Equilibria', in J. Persson and P. Ylikoski (*eds*), *Rethinking Explanation*, Dordrecht: Springer, pp. 149–62.

Kuorikoski, J. and Lehtinen, A. [2009]: 'Incredible Worlds, Credible Results', *Erkenntnis*, **70**, pp. 119–31.

Lehtinen, A. and Kuorikoski, J. [2007a]: 'Computing the Perfect Model: Why Do Economists Shun Simulation?', *Philosophy of Science*, **74**, pp. 304–29.

Lehtinen, A. and Kuorikoski, J. [2007b]: 'Unrealistic Assumptions in Rational Choice Theory', *Philosophy of the Social Sciences*, **37**, pp. 115–38.

Levins, R. [1966]: 'The Strategy of Model Building in Population Biology', *American Scientist*, **54**, pp. 421–31.

Levins, R. [1968]: *Evolution in Changing Environments; Some Theoretical Explorations*, Princeton, NJ: Princeton University Press.

Levins, R. [1993]: 'A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science', *Quarterly Review of Biology*, **68**, pp. 547–55.

Mäki, U. [1992]: 'On the Method of Isolation in Economics', in C. Dilworth (*ed*), *Intelligibility in Science*, Atlanta: Rodopi, pp. 319–54.

Mäki, U. [1994a]: 'Isolation, Idealization and Truth in Economics', in B. Hamminga and N. B. De Marchi (*eds*), *Idealization VI: Idealization in Economics*, Amsterdam: Rodopi, pp. 147–68.

Mäki, U. [1994b]: 'Reorienting the Assumptions Issue', in R. E. Backhouse (*ed*), *New Directions in Economic Methodology*, London: Routledge, pp. 236–56.

Mäki, U. [2000]: 'Kinds of Assumptions and their Truth: Shaking an Untwisted F-Twist', *Kyklos*, **53**, pp. 317–35.

Mäki, U. [2004]: 'Realism and the Nature of Theory: A Lesson from J H Von Thünen for Economists and Geographers', *Environment & Planning A*, **36**, pp. 1719–36.

Mäki, U. [2006]: 'Remarks on Models and Their Truth', *Storia del Pensiero Economico: Nuova Serie*, **3** (1), pp. 7–19.

Mäki, U. [2009]: 'Realistic Realism about Unrealistic Models', in H. Kincaid and D. Ross (*eds*), *Oxford Handbook of the Philosophy of Economics*, Oxford: Oxford University Press, pp. 68–98.

Mäki, U. [forthcoming]: 'Models and the Locus of their Truth', *Synthese*.

Mäki, U. and Marchionni, C. [2009]: 'On the Structure of Explanatory Unification: The Case of Geographical Economics', *Studies in History and Philosophy of Science*, **40**, pp. 185–95.

McMullin, E. [1985]: 'Galilean Idealization', *Studies in History and Philosophy of Science Part A*, **16**, pp. 247–73.

Melitz, J. [1965]: 'Friedman and Machlup on the Significance of Testing Economic Assumptions', *The Journal of Political Economy*, **73**, pp. 37–60.

Musgrave, A. [1981]: '"Unreal Assumptions" in Economic Theory: The F-Twist Untwisted', *Kyklos*, **34**, pp. 377–87.

Odenbaugh, J. [2005]: 'Idealized, Inaccurate but Successful: A Pragmatic Approach to Evaluating Models in Theoretical Ecology', *Biology and Philosophy*, **20**, pp. 231–55.

Orzack, S. H. and Sober, E. [1993]: 'A Critical Assessment of Levins's *The Strategy of Model Building in Population Biology* (1966)', *Quarterly Review of Biology*, **68**, pp. 533–46.

Ottaviano, G. I. P. [2003]: 'Regional Policy in the Global Economy: Insights from New Economic Geography', *Regional Studies*, **37**, pp. 665–73.

Ottaviano, G. I. P. [2007]: 'Models of "New Economic Geography": Factor Mobility vs. Vertical Linkages', in B. Fingleton (*ed*), *New Directions in Economic Geography*, Cheltenham: Edward Elgar, pp. 53–69.

Ottaviano, G. I. P. and Robert-Nicoud, F. [2006]: 'The "Genome" of NEG Models with Vertical Linkages: A Positive and Normative Synthesis', *Journal of Economic Geography*, **6**, pp. 113–39.

Ottaviano, G. I. P. and Thisse, J. [2004]: 'Agglomeration and Economic Geography', in J. Henderson and J. Thisse (*eds*), *Handbook of Urban and Regional Economics*, Amsterdam: Elsevier, pp. 2563–608.

Ottaviano, G. I. P., Tabuchi, T. and Thisse, J. [2002]: 'Agglomeration and Trade Revisited', *International Economic Review*, **43**, pp. 409–35.

Puga, D. [1999]: 'The Rise and Fall of Regional Inequalities', *European Economic Review*, **43**, pp. 303–34.

Rosenberg, A. [1992]: *Economics: Mathematical Politics or Science of Diminishing Returns?,* Chicago: University of Chicago Press.

Salmon, W. C. [1984]: *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

Sen, A. K. [1980]: 'Description as Choice', *Oxford Economic Papers*, **32**, pp. 353–69.

Sugden, R. [2000]: ‘Credible Worlds: The Status of Theoretical Models in Economics’, *Journal of Economic Methodology*, **7**, pp. 169–201.

Weisberg, M. [2006a]: ‘Robustness Analysis’, *Philosophy of Science*, **73**, pp. 730–42.

Weisberg, M. [2006b]: ‘Forty Years of “The Strategy”: Levins on Model Building and Idealization’, *Biology and Philosophy*, **21**, pp. 623–45.

Weisberg, M. and Reisman, K. [2008]: ‘The Robust Volterra Principle’, *Philosophy of Science*, **75**, pp. 106–31.

Wimsatt, W. C. [1980a]: ‘Randomness and Perceived Randomness in Evolutionary Biology’, *Synthese*, **43**, pp. 287–329.

Wimsatt, W. C. [1980b]: ‘Reductionistic Research Strategies and Biases in the Units of Selection Controversy’, in T. Nickles (*ed*), *Scientific Discovery*, Boston: Reidel, pp. 213–60.

Wimsatt, W. C. [1981]: ‘Robustness, Reliability and Overdetermination’, in M. B. Brewer and B. E. Collins (*eds*), *Scientific Inquiry and the Social Sciences*, San Francisco: Jossey-Bass, pp. 124–63.

Wimsatt, W. C. [1987]: ‘False Models as Means to Truer Theories’, in M. H. Nitecki and A. Hoffman (*eds*), *Neutral Models in Biology*, Oxford: Oxford University Press, pp. 33–55.

Wimsatt, W. C. [2007]: *Re-Engineering Philosophy for Limited Beings*, Cambridge, MA: Harvard University Press.

Woodward, J. [2006]: ‘Some Varieties of Robustness’, *Journal of Economic Methodology*, **13**, pp. 219–40.