ECONOMICS AS ROBUSTNESS ANALYSIS

Jaakko Kuorikoski, Aki Lehtinen and Caterina Marchionni

25.9. 2007

1. Introduction

Modern economic analysis consists largely in building abstract mathematical models and deriving familiar results from ever sparser modeling assumptions is considered as a theoretical contribution. Why do economists spend so much time and effort in deriving same old results from slightly different assumptions rather than trying to come up with new and exciting hypotheses? We claim that this is because the process of refining economic models is essentially a form of robustness analysis. The *robustness* of modeling results with respect to particular modeling assumptions, parameter values or initial conditions plays a crucial role for modeling in economics for two reasons. First, economic models are difficult to subject to straightforward empirical tests for various reasons. Second, the very nature of economic phenomena provides little hope of ever making the modeling assumptions completely realistic. Robustness analysis is therefore a natural methodological strategy for economists because economic models are based on various idealizations and abstractions which make at least some of their assumptions unrealistic (Wimsatt 1987; 1994a; 1994b; Mäki 2000; Weisberg 2006b). The importance of robustness considerations in economics ultimately forces us to reconsider many commonly held views on the function and logical structure of economic theory.

Given that much of economic research praxis can be characterized as robustness analysis, it is somewhat surprising that philosophers of economics have only recently become interested in robustness. William Wimsatt has extensively discussed robustness analysis, which he considers in general terms as *triangulation via independent ways of determination*. According to Wimsatt, fairly varied processes or activities count as ways of determination: measurement, observation, experimentation, mathematical derivation etc. all qualify. Many ostensibly different epistemic activities are thus classified as robustness analysis. In a recent paper, James Woodward (2006) distinguishes four notions of robustness. The first three are all species of robustness as similarity of the result under different forms of determination. *Inferential* robustness refers to the idea that there are different degrees to which inference from some given data may depend on various auxiliary assumptions, and *derivational* robustness to whether a given theoretical result depends on the different modelling assumptions. The difference between the two is that the former concerns derivation from data, and the latter derivation from a set of theoretical assumptions. *Measurement* robustness means triangulation of a quantity or a value by (causally) different means of measurement. Inferential, derivational and measurement robustness differ with respect to the method of determination and the goals of the corresponding robustness analysis. *Causal* robustness, on the other hand, is a categorically different notion because it concerns causal dependencies in the world, and it should not be confused with the epistemic notion of robustness under different ways of determination.

In Woodward's typology, the kind of theoretical model-refinement that is so common in economics constitutes a form of derivational robustness analysis. However, if Woodward (2006) and Nancy Cartwright (1991) are right in claiming that derivational robustness does not provide any epistemic credence to the conclusions, much of theoretical model-building in economics should be regarded as epistemically worthless. We take issue with this position by developing Wimsatt's (1981) account of robustness analysis as triangulation via independent ways of determination. Obviously, derivational robustness in economic models cannot be a matter of entirely independent ways of derivation, because the different models used to assess robustness usually share many assumptions. Independence of a *result* with respect to modelling assumptions nonetheless carries epistemic weight by supplying evidence that the result is not an artefact of particular idealizing modelling assumptions. We will argue that although robustness analysis, understood as systematic examination of derivational robustness, is not an empirical confirmation procedure in any straightforward sense, demonstrating that a modelling result is robust does carry epistemic weight by *guarding against error* and by *helping to assess the relative importance* of various parts of theoretical models (cf. Weisberg 2006b). While we agree with Woodward (2006) that arguments presented in favour of one kind of robustness do not automatically apply to other kinds of robustness, we think that the epistemic gain from robustness derives from similar considerations in many instances of different kinds of robustness.

In contrast to physics, economic theory itself does not tell which idealizations are truly fatal or crucial for the modeling result and which are not. Economists often proceed on a preliminary hypothesis or an intuitive hunch that there is some core causal mechanism that ought to be modeled realistically. Turning such intuitions into a tractable model requires making various unrealistic assumptions concerning other issues. Some of these assumptions are considered or hoped to be unimportant, again on intuitive grounds. Such assumptions have been examined in economic methodology using various closely related terms such as Musgrave's (1981) heuristic assumptions, Mäki's (2000) early step assumptions, Hindriks' (2006) tractability assumptions and Alexandrova's (2006) derivational facilitators. We will examine the relationship between such assumptions and robustness in economic model-building by way of discussing a case: geographical economics. We will show that an important way in which economists try to guard against errors in modeling is to see whether the model's conclusions remain the same if some auxiliary assumptions, which are hoped not to affect those conclusions, are changed. The case also demonstrates that although the epistemological functions of guarding against error and securing claims concerning the relative importance of various assumptions are somewhat different, they are often closely intertwined in the process of analyzing the robustness of some modeling results.

The structure of the paper is the following. Section 2 provides an introductory discussion of the notion of robustness. Section 3 explains why robustness is more important in economics than in some other sciences. Section 4 addresses the criticism that robustness is a non-empirical form of confirmation, and highlights the role of robustness in empirical testing. Section 5 presents our case. Section 6 addresses the criticisms levelled against robustness by discussing the independence of auxiliary assumptions. Section 6 further clarifies our account of derivational robustness by distinguishing it from other conceptions. Section 7 draws more general conclusions about the way in which the importance and rationale of robustness analysis should change our perception of the role of theory in economics. Section 8 concludes the paper.

2. Making sense of robustness

As Wimsatt uses the term, robustness means stability of a result under different and independent *forms of determination*, such as methods of observation, measurement, experiment or mathematical derivation. Robustness provides consilience via triangulation; a result is deemed real or reliable because a number of different and mutually independent routes lead to the same conclusion. If these separate ways of determination are indeed independent of each other, it would be a remarkable coincidence that they yield a similar conclusion if the conclusion does not correspond to something real.

Wimsatt refers to all methods of using various types of robustness considerations for distinguishing the real from the artifactual as *robustness analysis*. An important mode of robustness analysis is experimental triangulation. By performing experiments relying on different techniques and background theories, scientists can make sure that a putative phenomenon (as in Bogen and Woodward 1988) of interest is not merely an artefact of a particular experimental set-up. Similarly, different modes of measurement producing coherent results provide evidence for the existence of a single property that these different measurements are getting at. In philosophy of science, the multiple determination of Avogadro’s constant provides the most celebrated example of such an “argument from coincidence”. It exemplifies both these kinds of empirical robustness, and is taken to provide irrefutable grounds for the reality of molecules (Hacking 1983 54-55; Salmon 1984 214-220) .

Every experimental set-up and means of measurement has errors and biases. Sometimes we have prior knowledge of some of these problems, but there is always an element of residual uncertainty concerning the validity of an experiment or a measurement. *Independent ways of determining the same result reduce the probability of erroneous results due to errors or biases in individual ways of arriving at the result.* Fallible thinkers are better off avoiding long deductive chains because the chain as a whole is always weaker than its weakest link. In the case of multiple and independent ways of determination, the end-result is more secure than even the strongest individual reasoning. (Wimsatt 1987)

The required independence of the ways of determination can be spelled out using this error-avoidance viewpoint. In order for two or more ways of determination 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, its probability of failing to reach the correct result should not depend on whether the other one fails. Independence of errors therefore means that *given that the result holds* (or that it does not hold), the successes of the individual ways of determination in arriving at the result are independent of each other. Let DET_n be random variables expressing whether a method of determination n produces the result of interest R and RES a random variable expressing whether the result R actually holds or not. Independence of ways of determination can be defined using probabilistic independence as $DET_1 \perp DET_2 \ldots \perp DET_n | RES$.[1] If the methods of determination 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 R. (Bovens and Hartmann 2003, p. 96-97)

---

[1] 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 E1 is independent of another piece of evidence E2 (cf. Fitelson 2001).

3. Robustness in economics

Economics is not so much concerned with experimentation or measurement as with building analytically solvable mathematical models. These models are always based on various idealizations and abstractions, which make at least some of their assumptions unrealistic. Even a perfect economic model is idealized (Lehtinen and Kuorikoski forthcoming), since the point of economic models is to capture only the most important causal mechanisms and relationships. In physics, fundamental theories can in principle be used to determine how much distortion is introduced by each idealization (cf. Odenbaugh 2005; Weisberg 2006b). In contrast, in economics there is no such fundamental theory that tells the modeler which assumptions provide cause for alarm and which do not, and how one should go about making the models more realistic.

While economists cannot rely on theoretical frameworks for determining the importance of various idealizing assumptions, they may 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 which was previously neglected, providing links between variables that were already incorporated in the model, relaxing simplifying assumptions, restricting the domain of application (Levins 1993), specifying in more detail institutional or other contextual factors, and providing a more realistic account of individual behavior by allowing deviations from rationality or by allowing incomplete information. In some cases it is not clear what it would mean to replace a specific modeling assumption, such as a parameter value or a functional form, with a more realistic one.

Non-economists are often annoyed be the economists' seemingly sanguine attitude towards criticisms of unrealistic assumptions in their models: such criticisms are taken seriously, i.e. published in economic journals, only if they are accompanied with an alternate formal model that either shows how a modified assumption changes the conclusions of the previous model or shows that it does not change them. Economists obtain important information in both cases. If modifying an auxiliary assumption changes the result, the reason for the divergence of results can be traced to the difference between the original and the alternate model. If modifying the assumption does not change the result, this gives at least some assurance that the particular idealizations used in the two models are not crucially responsible for the results.

The existence of this practice (of requiring that a formal model be joined to criticisms of model-assumptions) provides evidence for our claim that economics *is* a form of robustness analysis. If mere unrealisticness of an assumption were sufficient to invalidate a model, it would be perfectly justifiable to accept verbal criticisms of assumptions. If it were easy to know which assumptions are realistic and which matter for the model-results, requiring a formal proof of non-robustness would be unnecessary.[2] The existence of this crucial methodological practice is thus premised on the importance of epistemic uncertainty concerning the consequences of unrealistic assumptions and the concomitant importance of robustness considerations and robustness analysis (see also Gibbard and Varian 1978).

When we say that economics is robustness analysis, we mean that model-building typically proceeds in such a way that some assumptions of an already existing model are modified in some respects. The modified model is often, though not necessarily, presented by a different individual economist. To be more precise, then, our claim is that *economics is collective derivational robustness analysis*.

A few comments are in order here. Michael Weisberg's (2006) account of robustness analysis contains four 'steps': 1) determine whether a set of models imply a common result 2) analyse whether this set has a common structure 3) formulate the robust theorem 4) conduct stability analysis to see what conditions defeat the connection between the common structure and the robust property. According to our account, there is no need for an individual economist to intentionally conduct robustness analysis. It is sufficient that someone presents an alternate model, and that this model retains some of the original assumptions. Furthermore, given that our factual claim concerns derivational robustness,

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

there is no need to require that stability analysis be conducted to justify calling the practice robustness analysis. Admittedly, when economists talk about robustness analysis, they often mean specific methods such as perturbation analysis or sensitivity analysis. Our definition clearly covers a broader range of activities than using these methods. We decline from providing neologisms for economists' practices, however, because we think that Woodward's term "derivational robustness" already captures its essence. Nancy Cartwright (2006) has argued that there is a grave need to conduct robustness analysis with respect to what she calls (after McMullin 1985) non-Galilean assumptions in economics models, namely those assumptions that do not participate in isolating a core causal mechanism, but that this is not done. If non-Galilean assumptions closely resemble tractability (or heuristic, or early step etc.) assumptions, we disagree about Cartwright's factual claim: economists do seem to be interested in changing tractability assumptions just as the Galilean assumptions. Our illustration, Geographical economics, is just one example. It is again important to bear in mind that what economists often call robustness analysis is not what we mean by this term.

4. The epistemic import of robustness analysis

The widespread 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; Sugden 2000). Empirical data is obviously the natural arbiter for evaluating the validity of theories. Even so, lacking access to the kind of data that would straightforwardly bear on whether a certain modelling result is accurate, it may be justifiable to use robustness analysis. Since the economic models are based on empirically dubious assumptions and some branches of economics have been criticized for the lack of empirical testing (Blaug 1980; Green and Shapiro 1994), the accusation of empirical emptiness of robustness is indeed more relevant in economics than in other fields.

According to Sugden, robustness analysis is a matter of comparisons between models, and thus does not licence 'the inductive leap' from models to economic reality. Similarly, Orzack and Sober (1993) 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' (p. 539). Here is how they formulate their argument:

> 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 (pp. 540-1).

Their claim is thus that it is the data rather than the robustness that does the confirming, or else the result is simply unfalsifiable and thus empirically empty. While we agree with them that robustness analysis does not by itself provide empirical confirmation, we think that their account provides a misleading view of the importance and role of robustness for empirical tests.

In our view, Orzack and Sober's example actually shows how robustness considerations are crucial in determining what can be confirmed with given data. Their argument rests on a crucial 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 and think that a theory or a model is confirmed or disconfirmed only as a whole – the only question is whether the model fits the data. But an additional inferential step is needed to claim that a model says anything interesting about the phenomenon on the basis of its fit to a set of data. If two models are actually in competition but cannot be distinguished empirically, 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 models is confirmed. Confirmation may concern the assumptions of the model, or the consequences of those assumptions. We are often interested in knowing that the causally important parts of our models, rather than the unimportant auxiliary assumptions, are confirmed or disconfirmed. Testability of parts of models requires robustness because, unless we know which conclusions are common to the two models, it is impossible to say which parts of the models are supported by the data. 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 uncertain assumptions and is relevant precisely when there is no direct way of empirically ascertaining whether the conclusions of our inferences are true.

Now suppose that two models M_1 =($C\&V_1\&V_3$) and M_2=($C\&V_2\&V_3$), consisting of a common theoretical element C and variable auxiliary assumptions V_i, have a common consequence R_M, and that given their differences, they also have consequences not common to the two models, R_{M1} and R_{M2}, respectively. Initially there is a set of data E that supports the common consequence R_M. The models can be assumed to be empirically equivalent with respect to this consequence. Suppose that we then obtain new evidence E_1 which is consistent with R_M, and R_{M1} but not with R_{M2}. Then we have learned that the data supports $C\&V_1\&V_3$ but not $C\&V_2\&V_3$. Because of the common and empirically supported implication, we now have a prima facie reason to suspect that there is something wrong specifically with V_2. Since we could not have made this inference without there being nonrobust implications, testability requires nonrobust theorems, just as Orzack and Sober argue. However, it also requires robust elements. To see this, consider a scenario in which we start with M_1 =($C\&V_1\&V_3$) and M_2=($C\&V_2\&V_3$), and we know that R_{M1} is derivable from M_1 and that R_{M2} is derivable from M_2, but we do not know whether R_{M1} is derivable from M_2 and whether R_{M2} is derivable from M_1. Furthermore, empirical evidence E supports both R_{M1} and R_{M2}. If we then find out that R_{M1} is not consistent with M_2, M_2 is falsified (since it cannot be consistent with E). The question, then, is what we can say about the assumptions in the two models without finding out whether R_{M2} is a common consequence of the two models. Only if we find that R_{M2} is consistent also with M_1 can we say that V_2 was the culprit. Testability thus requires both robust and nonrobust elements.

Another vocal critic of the evidential value of robustness is Nancy Cartwright. She argues (Cartwright 1991) that if different econometricians use different functional forms for deriving the same results, this proves nothing about the truth of the theorems derived. She argues that if we look at thirty different functional forms, but 'God's function is number thirty-one, the first thirty do not teach us anything'. Cartwright's criticism is based on the idea that a perfect model exists, one that provides the-whole-truth and nothing-but-the-truth (cf. Teller 2001). This must be so because otherwise it would not be possible to imagine that only one functional form describes reality. In physics, the idea of the one true functional form might be plausible in some modeling situations, but economics deals with complex phenomena within open systems characterized by heterogeneity and transient parameter values. For many auxiliary assumptions, there simply is no unique true formulation or value.

Cartwright argues that the inference based on robustness analysis does indeed lend credence to a robust theorem if we have some independent knowledge that the different functional forms have some common properties. Yet, if we already know that some functional form should have a certain property and that the result depends on this property, what is the point of doing robustness analysis? It is precisely when we lack such knowledge that robustness analysis provides us epistemically relevant information. The whole point of robustness analysis is to determine which factors are really important and which not. Thus, to show that a result can be derived with two or more functional forms gives us assurance that what was believed to be irrelevant to the question under study – in this case factors encoded in the different functional forms – is indeed irrelevant. What is important for robustness analysis is not which functional form, if any, is the true one, but rather that the truth or falsity of the different functional forms does not matter. As Levins puts it, the point of robustness analysis is to determine the extent we can get away with *not* knowing all the details and still understand the system (Levins 1993, 554).

Levins' original account makes clear that the point with robustness lies precisely in the idea that all models are strictly speaking false, if falsity means the absence of the-whole-truth and nothing-but-the-truth.[3] Once we discard the received view of theories as axiomatic logical systems and switch to speaking more generally about models as more or less accurate representations, it is easy to see that the notions of truth or falsity of a model provide merely the extreme points for (at least) two dimensions of partial truth: comprehensiveness and accuracy of representations. Talk of truth or falsity simpliciter does not make much sense in the world of models. For our purposes, it is not necessary to provide an extensive account of partial truth. All we need is the idea that models as well as their individual assumptions may approximate truth to various degrees.[4] The usual case is one in which we know that all the models are false in the sense that they contain various idealisations and abstractions and are thus not completely true. Yet they all might contain a grain of truth, and we do not know which combination of auxiliary assumptions provides a model that is closest to the truth. More importantly, before conducting robustness analysis, we do not know whether a result depends on the different auxiliary assumptions or not: we do not know the epistemic status of the theorem. When we discover that a result is implied by multiple models, we can be more confident that it does not depend on the idealisations we have made (2006a). Although it does not provide a means of *finding* causally important mechanisms, robustness analysis serves to distinguish the causally important mechanisms from the 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 a model fails to depict the important factors, it will not give a truthful view of the situation even if its results are robust with respect to many auxiliary assumptions. The contribution that robustness analysis provides is that it allows the modeler to be just a little bit more certain that what he or she hoped is irrelevant is in fact irrelevant.

[3] This distinction was proposed in Sen (1980). See Mäki (1992; 1994a) for further analysis.
[4] The failure of the framework provided by French and da Costa (2003) to do this in a reasonable way is the reason we are not using it here (see Pincock 2005) .

5. An illustration: geographical economics models

Geographical economics (henceforth GeoEcon) is a recent approach to spatial issues developed within economics that endeavours to explain the spatial location of economic activity. Krugman (1991) provides the first and the core model.[5] Following its appearance, a growing theoretical literature refines and extends this original stylized model. The CP model and its close follow-ups (summarized in Fujita, Krugman and Venables1999) depends crucially on a set of unrealistic assumptions, or "modeling tricks", as geographical economists sometimes refer to them. 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 identified as robustness analysis: many of the subsequent GeoEcon models serve the function of checking whether the main conclusions of the CP model remain valid when some of the unrealistic assumptions are altered. We begin by briefly setting out the main elements of the core model, and then look at a few GeoEcon models that explore the robustness of its results.

Krugman (1991) employs the Dixit-Stiglitz (1977) general equilibrium model of monopolistic competition with transportation costs and labor mobility to derive a core-periphery pattern, that is, a situation in which the bulk of economic activity is located in one region. In particular, the model makes the following assumptions: There are two regions, identical in all respects, and two sectors in the economy. The perfectly competitive (agricultural) sector employs unskilled labor, which is distributed equally between the two regions and cannot move across regions. The monopolistically competitive (manufacturing) sector uses only one input, skilled labor, which can move across regions, and 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. The trade of the manufacturing good produced is subject to transportation costs. In order to avoid modeling a separate sector,

[5] Krugman's (1991) model is often referred to as the CP model (C for 'core' and P for 'periphery').

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. This can be seen as the core causal mechanism. The centripetal force is constituted be the circular causation of forward and backward linkages: because of increasing returns and transportation costs, firms have an incentive to locate in the larger market for their product in order to economize on transportation costs, and workers/consumers have an incentive to locate where the producers of manufacturing goods are, in order to economize on living costs and to benefit from variety. This creates a process of circular causation: the more firms and workers there are in the region, the more attractive the region becomes for further firms and workers (market-size effect). The centripetal force is counteracted by the need to serve the immobile factor which remains equally distributed across the two regions and by market-crowding effects: in the larger region, firms face higher competition and higher input prices.

The result [R] of the model is that agglomeration is more likely to occur in sectors where increasing returns are intense, market power is strong, and transportation costs are low. Lower transportation costs, strong market power and intense increasing returns in fact reduce the effects of market crowding, and hence of the centrifugal force vis-à-vis the centripetal force. Importantly, the CP model shows that changes of transportation costs affect the balance between centripetal and centrifugal forces in non-linear way (seeOttaviano 2003). [6]

[6]The CP model also displays two important qualitative features: (i) catastrophic agglomeration, that is, 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, that is, transitory shocks can have permanent effects. These two features are generally regarded as distinctive of the GeoEcon approach. But Puga (1999) shows that the existence of congestion in the agglomerating region can render the transition from dispersion to agglomeration gradual, and the same happens if some heterogeneity were introduced across firms. Ottaviano et al. (2002) predict catastrophic agglomeration without hysteresis. In these cases however an economic explanation for the different results is given.

The core model relies on very specific functional forms for utility (namely, CES functions) and production functions (a homothetic function where skilled labor is the only input), and iceberg transportation technology. These assumptions are made mainly for convenience or mathematical tractability, and it is hoped that the results [R] do not crucially depend on this specific set of assumptions.[7]

Later models explore the robustness of the properties and results of the CP model. In particular, Ottaviano, Tabuchi and Thisse (2002) employ quadratic sub-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 that vary with prices, and rather than being fixed by a markup rule (as in the CP model), equilibrium prices depend on the fundamentals of the market. The main conclusions of the CP model, however, are found to be robust in spite of these changes. Ottaviano et al. (2002, p. 432) conclude that "The main results in the literature do not depend on the specific modeling 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), labeled Footloose Entrepreneurs (FE), derive a core-periphery pattern driven by mobility of labor, as in the CP model, but assume a different specification of the production function. Whereas the core model uses a homothetic function in which only skilled labor appears, FE models use the Flam and Helpman (1987) functional form, which assumes that firms use both skilled and unskilled labor. This change has a significant implication. In the core model, free entry determines the size of the firms, and hence the number of firms functions as the adjustment variable. In the FE models, the number of firms varies with wage instead. The FE models nonetheless derive the same qualitative results as the CP models: agglomeration is positively affected by the intensity

[7] Notice that even with these simplifications, deriving the results of the core model still requires resorting to numerical computations.

of the increasing returns, and negatively affected by the level of transportation costs and market power.[8]

The models whose results are found to be robust share a set of substantive assumptions. Despite modifications to some components of the core model, a set of ingredients remain invariant across these alternative frameworks: the presence of firm-specific economies of scale, imperfect competition, transportation costs and labor mobility, giving rise to the centrifugal and centripetal forces. The theoretical claim of the GeoEcon models is that it is this common set of substantive assumptions, the core causal mechanism, that is responsible for the same result [R] derived across different models. A series of recent works demonstrate that the main models of GeoEcon are isomorphic at equilibrium in a meaningful state space, and this purportedly explains why, despite differences in their functional specifications, they obtain the same qualitative results. Robert-Nicoud (2005) shows this to be the case for models assuming labor mobility, namely the CP model (Krugman 1991) and FE model (Forslid and Ottaviano 2003).

To understand the role of robustness analysis, it is useful to look at modifications of the models where robustness breaks down. For example, Ottaviano and Thisse (2004) show that more realistic models that include an additional spatial cost (for example, congestion or transport costs for the good of the agricultural sector) demonstrate that the value of the transport cost relative to this additional spatial cost is crucial for determining whether dispersion or agglomeration occurs. For instance, adding transport costs for the agricultural good changes the qualitative results [R] of the CP model. In such model, when the transport cost for the agricultural good is low, agglomeration occurs for intermediate levels of the transport cost of the manufacturing good, as in the CP model. But, in contrast to the CP model, when the cost of transporting the agricultural good is high, industry is always dispersed for high levels of the cost of transporting the manufacturing good.

[8] The core model states that: $C(x(i) = w[F + \rho x(i)]$, where x(i) is the firm output, F is the fixed labour requirement, ρ is the variable labour requirement, and *w* is the skilled labour wage. The Flam and Helpman functional form is: $C(x(i) = wF + w_u \rho x(i)$, where w_u is the unskilled labour wage rate.

However, the divergence of results is not taken to invalidate the theoretical claim connecting the core mechanism of the CP model to [R].[9] The reason is that there is an economic explanation for the divergent results. First, dispersion (and not agglomeration) now occurs for low transportation costs of the manufacturing good because the price differential of the agricultural good constitutes a further centrifugal force that was absent in the CP model. More importantly, for high transportation costs, the centrifugal force (the crowding effect on the market for the manufacturing good) is exactly the same as in the CP model. This confirms that, in the absence of a significant additional cost, the relationship between high transport costs of the manufacturing good and dispersion is as predicted in the CP model. Second, the result 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). This finding is consistent with the claim that robustness analysis serves to track the relative importance of various parts of theoretical models, and to discriminate between potentially relevant causal factors from causally irrelevant ones.

The breakdown of robust conclusions does not constitute a problem as long as it can be given a plausible economic interpretation. Whether this economic interpretation is correct is an empirical matter, and so it should be ultimately determined. As claimed above, robustness analysis does not constitute a form of empirical validation. The important point, however, is that if there is no economic explanation for the break-down of the robust conclusions, there is little reason to confront the models with empirical evidence; they are regarded as epistemically worthless.

To see more precisely the difference between the case in which the results changed, and the one in which they did not, it is useful to distinguish between two categories of

[9] U-shaped relationship between transport costs and agglomeration versus Ω-type relationship. The bell-shape relationship also occurs when there are input-output linkages and labor is immobile; when there are urban crowding costs; or when there is heterogeneity in migration behavior. The bell-shaped relationship also corresponds to a stylized fact (the bell-shaped curve of spatial development): a core-periphery pattern emerges in early stages of development, but as development proceeds, the core and the periphery (urban-rural areas) converge.

auxiliary assumptions: idealizing assumptions about factors known or presumed to have an impact on the phenomenon (to be relaxed at later stages), and very specific modeling assumptions that are made for mathematical convenience and are hoped to be innocuous or irrelevant for the economic result of the model (tractability assumptions or derivation facilitators). The practice of model building in economics seems to be such that assumptions of the first kind are relaxed for the sake of increased realisticness, increased applicability of the model, or better predictions. The breakdown of robust results in this case confirms the belief that the factor in question does have an impact on the phenomenon, and the model is used to study how the introduction of the new factor affects the working of the mechanism and hence the conclusions of the model. Tractability assumptions instead are typically replaced with different unrealistic assumptions, that is, assumptions that are false in a different way. The failure of robustness with respect to tractability assumptions is problematic because it shows that the result is an artifact of the specific set of tractability assumptions. Instead, if a result turns out to be robust, the dependency between the set of substantive economic assumptions and the result gets an extra hedge in terms of credibility. This is what we have identified as the function of guarding against errors: robustness analysis guards against unknown consequences of assumptions that are hoped to be innocuous. The different roles played by alternative specifications of functional forms and technologies in GeoEcon exemplify this practice and its function very clearly.

6. Independence of derivations

Robustness confers credibility to a result if the ways of determination with respect to which the result is robust are independent from each other. Thus, as Orzack and Sober (1993) have pointed out, the intelligibility of the concept of robustness crucially depends on the possibility of giving an account of the independence of the ways of determination.

Existing discussion on the independence in robustness analysis seems to take it for granted that independence is a matter of independence *between models* (see Orzack and

Sober 1993). What does independence between models mean? To be sure, it cannot mean that any two models taken to be independent do not share any assumptions or variables. If it did, it would be very difficult to find independent models. Orzack and Sober discuss what they call *logical independence,* which has to do with the truth values of models, and *statistical* independence. The problem with these suggestions is that they do not seem to be very well suited for analysing models (in the sense in which economists talk about models, not in the sense of how they are understood in the semantic view); models are very seldom logically independent from each other because they do not have clearly defined truth-values in the first place, or are always false, and assigning probabilities to the truth of models does not seem reasonable either. However, Orzack and Sober are right about the fact that Levins' notion of robustness does not really incorporate the independence assumption. When Levins discusses robustness, he refers to the idea that robust theorems are those that *share* the *biological* assumptions (1966; 1968). Beyond noting that the various models share the biological assumptions, he does not say very 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 full-fledged theory. Something similar is going on in economic models. As exemplified by the GeoEcon case, different economic models share some substantive economic assumptions, possibly those that specify the core mechanism, but not all.

Let us express Levins' conception of robustness as follows:[10] Let $C\&V_i$ denote a model M_i that is based on combining the common part of the set of models with the specific part V_i. Let $A \vdash B$ denote 'B is derivable from A' (within some standard formal system). Robustness of the relationship $C \vdash R_M$ requires that

$(C\&V_1) \vdash R_M$, and

[10] Levins (1993) proposes to evaluate robustness with the following framework. Let R_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

$$CV_1+CV_2+\ldots=C(V_1+V_2+\ldots)\rightarrow R_M,$$

and if the set of V_i's is exhaustive, then $(V1+V_2+\ldots)=1$, and $C\rightarrow R_M$. Since Levins does not define the addition sign (+) in this formalism, it is difficult to tell what exactly is being claimed.

$(C\&V_2) \vdash R_M$, and...

$(C\&V_n) \vdash R_M$.

Robustness of the derivation is hardly interesting if R_M can be derived without making any auxiliary assumptions, it is therefore natural to think that $\sim(C \vdash R_M)$, although this is not a formal requirement of robustness. Levins' treatment of robustness gives the impression that the set of auxiliary assumptions must be exhaustive, but we think this is too restrictive. In practice the typical case is one in which it is not possible to define an exhaustive set of possible auxiliary assumptions. Allowing for a non-exhaustive set weakens the degree to with which each derivation, which use a different set of auxiliary assumptions, shows that the relationship is robust, but it does not remove the epistemic relevance of such derivations altogether.

It is important to realize that even though the various models in M are not independent from each other because they share some assumptions, it is the *independence of individual assumptions* in models that is really crucial for derivational robustness. As Wimsatt (1980, p. 310) notes, the models must be similar so that we can compare them, and isolate their similarities and differences. What a robust theorem claims is that the result R_M depends on some central mechanism C, irrespective of the details of the other assumptions in the models. Levins' (1966, p. 126) unclear but intuitive claim that 'our truth is the intersection of independent lies' can be taken to mean that the fact that result R_M can be derived from mechanism C does not depend on the various auxiliary assumptions used in the derivation. Various falsities involved in the different possible derivations do not matter because we know that result R_M does not depend on the falsehoods. In other words, since the result R_M can be derived with all possible falsehoods $V_1,...,V_n$, their falsehood does not matter. In the GeoEcon case, we saw that changing the specifications of functional forms and technologies, while retaining the assumptions depicting the core causal mechanism, yielded the same qualitative conclusions. The derivation of [R] from different sets of unrealistic assumptions increases the confidence that [R] depends on the core economic mechanism rather than on specific unrealistic modeling choices.

Our discussion has so far centered on the role of robustness analysis in assessing the relative importance of modeling assumptions. The question that remains is whether Levinsian robustness analysis can be conceived as a species of general robustness analysis and whether the same epistemic rationale applies to it. What is needed is an account of how the different auxiliary modeling assumptions ($V_1 \ldots V_n$) could be thought of as not sharing the same biases and other sources of possible error. In general, robustness required that the success of ways of determination be independent given knowledge of whether the result holds ($DET_1 \perp DET_2 \ldots \perp DET_n | RES$). If we plug in the models, we get $((C\&V_1) \vdash R_M \perp (C\&V_2) \vdash R_M \ldots \perp (C\&V_n) \vdash R_M | R_M)$. The problem is that $((C\&V_n) \vdash R_M)$ are not random variables. Talk of reliability, biases and errors seems out of place here since we are dealing with logical relations of formal derivability, not causal processes of measurement or experimentation. Whether a relation of derivability holds cannot be a matter of probabilities.

Our proposal for solving this problem is to go subjectivist and relativise the epistemic gain from robustness to the epistemic situation of the modeler or the relevant scientific community at a certain time. Modeling is an act of inference from a set of substantial assumptions to a conclusion. However, this process of inference usually requires additional auxiliary assumptions to be feasible. These auxiliary assumptions can lead the process of inference astray in the sense that it leads to a conclusion that is contrary to the substantial assumptions. In this sense, auxiliary assumptions can induce errors in the modeling process. For robustness considerations to be meaningful in the context of formal modeling, it is sufficient that the person deriving robust theorems does not have reasons to believe that the auxiliary assumptions will systematically lead astray. Similarly, independence means that the modeler should have no positive reasons to believe that if one of the auxiliary assumptions induces a significant error in the result, so does another one. Given that the modeling result of interest (R_M) is correct, the prior probabilities concerning whether R_M can be derived from $C\&V_1$ or $C\&V_2 \ldots C\&V_n$ should be independent. In the GeoEcon case, if assuming CES or quadratic utility functions had exactly the same mathematical implications, then they would not be independent in that they would induce the same kind of falsehood to the model. Instead, CES utility functions and sub-quadratic utility functions are known to be independent at least in the weak sense that neither is a special case of the other. Quadratic utility functions, for instance, entail that demand and substitution elasticities vary with prices (in the CES variant, they do not) and equilibrium prices depend on the fundamentals of the market (in the CES variant, they do not).

In sum, economic models in a given sub-field are often far from independent because they usually differ only with respect to a few particular modeling assumptions. Robustness analysis in economics is thus usually a special, degenerate form of general robustness analysis in Wimsatt's sense: testing the robustness of a result with respect to a particular modeling assumption which is usually obviously unrealistic. If a modeling result is robust with respect to particular modeling assumptions, the empirical falsity of these particular assumptions does not provide grounds for criticizing the result. As Kent Staley (2004) puts it, the evidence provided by the model is more secure against false assumptions. Thus analytical robustness can provide epistemic credence for the reality of an economic phenomenon corresponding to the modeling result, at least in this purely negative manner. In the analogous case of robustness of an empirical result under independent experimental procedures, the support is stronger if we know or have *positive* reasons to believe that the procedures are causally or statistically independent in the sense that they do not share the same characteristic errors. If we merely have no particular positive reasons to believe them to be dependent, the credibility conferred by different ways of determination is obviously weaker, but not non-existent.

Orzack and Sober are of course right in claiming that purely analytic robustness analysis itself does not provide empirical confirmation, since no new observations are made. However, the value of analytic robustness analysis lies in rooting out error from our inferences from diverse and uncertain assumptions to conclusions. If these assumptions have no empirical merit in the first place, then clearly robustness has no epistemic value – it is just a similarity relation between members of a peculiar set of abstract entities. But in case of economics, this would amount to claiming that economic modeling assumptions did not have any empirical content in the first place. Although this is a remotely plausible view that may even have been entertained, it is clearly a rather extreme position.

7. Economics as a Babylonian science

Viewing economic model building as a form of robustness analysis dramatically changes our perception of the status and role of economic theory. Over the decades, the self-image of economics has been consciously molded to fit the template of a Millian deductive science in the sense that economic theorizing is explicitly based on fundamental economic axioms and logical derivation of empirical results from these axioms. Wimsatt calls this particular ideal of science Euclidian (1981). However, the former discussion shows that one of the main epistemic strategies of economics seems to be incoherent with this image; a situation Rosenberg calls the deepest metaphysical mystery of economics (Rosenberg 1995). Little interest is shown in empirically securing the foundations of the theory to the maximum extent, which would be the obvious way for a true Euclidian to proceed. Indeed, economic model building is notoriously insensitive to repeated empirical refutations of the rationality axioms supposedly providing the basis for the whole enterprise. Instead, enormous intellectual effort is poured into analytical derivations of often familiar results from slightly different assumptions. Rather than a stubborn metaphysical prejudice, this epistemic strategy is best seen as a form of robustness analysis, which aims at isolating and providing epistemic credence to economic phenomena corresponding to the robust results or theorems, not direct hypothetico-deductive support for a grand axiomatic theory as such. This strategy fits better with another image of science, one that Richard Feynman (1992 [1965], p. 46) and Wimsatt (Wimsatt 1981)have called Babylonian.

Economics is a Babylonian science, because it deals with complex phenomena resulting from underlying diversity, not uniformity. For most economic phenomena of interest, there simply is no single true functional form decreed by God. The rationality postulates do not describe invariant natural constants that would provide secure foundations for modelling if we only could measure them accurately. If there were such invariants, the best way to increase reliability would be to eliminate all possible errors concerning the axioms describing the invariants: if there are no errors in the axioms, there will be no errors in the whole deductive edifice (cf. Rosenberg 1992 ,chpt. 5). However, since real economic systems are fundamentally heterogeneous and open, error cannot be so eliminated. To increase reliability, the next best way is to control for the effects of inevitable errors. The Babylonian character of economics thus explains the fact that none of the so called fundamental axioms of economics is truly sacrosanct (cf. Hausman 1992, 52). It also makes it easy to understand why new branches of economics typically start from a special case which is then generalised and modified (Balzer and Dreier 1999).

In order to control for the epistemic consequences of the inevitable errors in modelling assumptions, the theoretician is free to violate any axioms she pleases, as long as she is doing it in a way that helps to locate sources of possible error relevant for the result of interest, i.e. with an explicit formal model showing how the result of a model depends on a particular problematic modelling assumption. This epistemic groping is necessary, since economic theory does not in itself provide much guidance on what kind of idealizations are fatal to certain empirical conclusions and which are not.

Although many argue for the economic approach on the basis of its unificatory credentials, the importance of robustness and the Babylonian nature of economics entail that substantial unification actually *increases* the susceptibility of the theoretical system to inevitable errors. If diverse theorems and sub-models are really substantially dependent on a small set of unifying axioms, the ramifications of errors in the axioms are more destructive and far reaching (Wimsatt 1981). Instead, the proper function of the theory of individual utility maximization in its various guises is to be understood with what Boyd and Richerson (1987) call a *generalized sample model*. The generalized sample model is a generalization over more detailed and context sensitive models which are still relatively simple, but are nonetheless explanatory and perhaps predictive. The role of the generalized sample theory is not to directly tell the truth about the deepest unified nature of the variety of more detailed models (as in physics), but rather to help organize cognitive work and facilitate communication between specialists. The role of the theory is to help understand models, not the world.

8. Conclusions

The practice of economic theorizing largely consists of building models with slightly different assumptions yielding familiar results. We have argued that this practice makes sense when seen as derivational robustness analysis. Robustness analysis is a sensible epistemic strategy in situations where we know that our assumptions and inferences are fallible, but not in what situations and in what way. Derivational robustness analysis guards against errors in theorizing when the problematic parts of the ways of determination, i.e. models, are independent of each other. In economics in particular, proving robust theorems from different models with diverse unrealistic assumptions helps us to evaluate what results correspond to important economic phenomena and what are merely artefacts of particular auxiliary assumptions. We have addressed Orzack and Sober's criticism against robustness as an epistemically relevant feature by showing that their formulation of the epistemic situation in which robustness analysis is useful is misleading. We have also shown that their argument actually shows how robustness considerations are necessary for evaluating what a given piece of data can support. We have also responded to Cartwright's criticism by showing that it relies on an untenable hope of a completely true economic model.

Viewing economic model building as robustness analysis also helps to make sense of the role of the rationality axioms that apparently provide the basis of the whole enterprise. Instead of the traditional Euclidian view of the structure of economic theory, we propose that economics should be approached as a Babylonian science, where the epistemically secure parts are the robust theorems and the axioms only form what Boyd and Richerson call a generalized sample theory, whose the role is to help organize further modelling work and facilitate communication between specialists.

References

Alexandrova, Anna (2006): "Connecting Economic Models to the Real World: Game Theory and the FCC Spectrum Auctions", *Philosophy of the Social Sciences,* vol. 36, no. 2, pp. 173-192.

Balzer, Wolfgang & Dreier, Volker (1999): "The Structure of the Spatial Theory of Elections", *British Journal for the Philosophy of Science,* vol. 50, no. 26, pp. 613-638.

Blaug, Mark (1980): *The methodology of economics : or, How economists explain*, Cambridge University Press, Cambridge Eng. ; New York.

Bogen, James & Woodward, James (1988): "Saving the Phenomena", *The Philosophical Review,* vol. 97, no. 3, pp. 303-352.

Bovens, Luc & Hartmann, Stephan (2003): *Bayesian epistemology,* Clarendon Press, Oxford.

Boyd, Robert & Richerson, Peter (1987): "Simple Models of Complex Phenomena: The Case of Cultural Evolution", *in The Latest on the Best*, ed. J. Dupré, MIT Press, Cambridge, MA/London, pp. 27-52.

Cartwright, Nancy (2006): "The Vanity of Rigour in Economics: Theoretical Models and Galilean Experiments", *in The 'experiment' in the history of economics*, eds. P. Fontaine & R. Leonard, Routledge, London, pp. 135-153.

Cartwright, Nancy (1991): "Replicability, Reproducibility, and Robustness: Comments on Harry Collins", *History of Political Economy,* vol. 23, no. 1, pp. 143-155.

Costa, Newton C. A. d. & French, Steven (2003): *Science and partial truth : a unitary approach to models and scientific reasoning*, Oxford University Press, New York.

Dixit, Avinash K. & Stiglitz, Joseph E. (1977): "Monopolistic Competition and Optimum Product Diversity", *The American Economic Review,* vol. 67, no. 3, pp. 297-308.

Feynman, Richard P. (1992 [1965]): *The character of physical law*, Penguin Books, London.

Fitelson, Branden (2001): "A Bayesian Account of Independent Evidence with Applications", *Philosophy of Science*, vol. 68, no. 3.

Flam, Harry & Helpman, Elhanan (1987): "Vertical Product Differentiation and North-South Trade", *The American Economic Review,* vol. 77, no. 5, pp. 810-822.

Forslid, Rikard & Ottaviano, Gianmarco I. P. (2003): "An Analytically Solvable Core-Periphery Model", *Journal of Economic Geography,* vol. 3, no. 3, pp. 229-240.

Fujita, Masahisa, Krugman, Paul & Venables, Anthony (1999): *The spatial economy : cities, regions, and international trade*, MIT Press, Cambridge, Mass.

Gibbard, Allan F. & Varian, Hal R. (1978): "Economic Models", *Journal of Philosophy,* vol. 75, no. 11, pp. 664-677.

Green, Donald P. & Shapiro, Ian (1994): *Pathologies of rational choice theory : a critique of applications in political science*, Yale University Press, New Haven.

Hacking, Ian (1983): *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*.

Hausman, Daniel M. (1992): *The inexact and separate science of economics,* Cambridge University Press, Cambridge ; New York.

Hindriks, Frank A. (2006): "Tractability Assumptions and the Musgrave-Mäki Typology", *Journal of Economic Methodology,* vol. 13, no. 4, pp. 401-423.

Krugman, Paul (1991): "Increasing Returns and Economic Geography", *The Journal of Political Economy,* vol. 99, no. 3, pp. 483-499.

Lehtinen, Aki & Kuorikoski, Jaakko (forthcoming): "Computing the Perfect Model: Why do Economists Shun Simulation?", *Philosophy of Science,* vol. 74, no. 3.

Lehtinen, Aki & Kuorikoski, Jaakko (2007): "Unrealistic Assumptions in Rational Choice Theory", *Philosophy of the Social Sciences,* vol. 37, no. 2, pp. 115-138.

Levins, Richard (1993): "A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science", *Quarterly Review of Biology,* vol. 68, no. 4, pp. 547-555.

Levins, Richard (1968): *Evolution in changing environments; some theoretical explorations*, Princeton University Press, Princeton, N.J.

Levins, Richard (1966): "The Strategy of Models Building in Population Biology", *American Scientist,* vol. 54, pp. 421-431.

Mäki, Uskali (2000): "Kinds of Assumptions and their Truth: Shaking an Untwisted F-Twist", *Kyklos,* vol. 53, no. 3, pp. 317-335.

Mäki, Uskali (1994a): "Isolation, Idealization and Truth in Economics", *in Idealization VI: Idealization in economics*, eds. B. Hamminga & N.B. De Marchi, Rodopi, , pp. 147-168.

Mäki, Uskali (1994b): "Reorienting the Assumptions Issue", *in New directions in economic methodology*, ed. R.E. Backhouse, Routledge, , pp. 236-256.

Mäki, Uskali (1992): "On the Method of Isolation in Economics", *in Intelligibility in Science*, ed. C. Dilworth, Rodopi, Atlanta and Amsterdam, pp. 319-354.

McMullin, Ernan (1985): "Galilean Idealization", *Studies In History and Philosophy of Science Part A*, vol. 16, no. 3, pp. 247-273.

Musgrave, Alan (1981): "'Unreal Assumptions' in Economic Theory: The F-Twist Untwisted", *Kyklos*, vol. 34, no. 3, pp. 377-387.

Odenbaugh, Jay (2005): "Idealized, Inaccurate but Successful: A Pragmatic Approach to Evaluating Models in Theoretical Ecology", *Biology and Philosophy*, vol. 20, no. 2-3, pp. 231-255.

Orzack, Steven H. & Sober, Elliott (1993): "A Critical Assessment of Levins's the Strategy of Model Building in Population Biology (1966)", *Quarterly Review of Biology*, vol. 68, no. 4, pp. 533-546.

Ottaviano, Gianmarco I. P. (2007): "Models of 'New Economic Geography': Factor Mobility Vs. Vertical Linkages", *in New directions in economic geography*, ed. B. Fingleton, Edward Elgar, Cheltenham (UK), pp. 53.

Ottaviano, Gianmarco I. P. (2003): "Regional Policy in the Global Economy: Insights from New Economic Geography", *Regional Studies*, vol. 37, no. 6-7, pp. 665-673.

Ottaviano, Gianmarco I. P. & Robert-Nicoud, Frederic (2006): "The 'Genome' of NEG Models with Vertical Linkages: A Positive and Normative Synthesis", *Journal of Economic Geography*, vol. 6, no. 2, pp. 113-139.

Ottaviano, Gianmarco I. P. & Thisse, Jacques-Francois (2004): "Agglomeration and Economic Geography", *in Handbook of Urban and Regional Economics*, eds. J. Henderson & J. Thisse, Elsevier, Amsterdam.

Pincock, Christopher (2005): "Overextending Partial Structures: Idealization and Abstraction", *Philosophy of Science*, vol. 72, pp. 1248-1259.

Puga, Diego (1999): "The Rise and Fall of Regional Inequalities", *European Economic Review*, vol. 43, no. 2, pp. 303-334.

Robert-Nicoud, Frederic (2005): "The Structure of Simple 'New Economic Geography' Models (Or, on Identical Twins)", *Journal of Economic Geography*, vol. 5, no. 2, pp. 201-234.

Rosenberg, Alexander (1995): "The Metaphysics of Microeconomics", *Monist*, vol. 78, no. 3, pp. 352.

Rosenberg, Alexander (1992): *Economics : mathematical politics or science of diminishing returns?*, University of Chicago Press, Chicago.

Salmon, Wesley C. (1984): *Scientific explanation and the causal structure of the world,* Princeton University Press, Princeton, N.J.

Sen, Amartya K. (1980): "Description as Choice", *Oxford Economic Papers,* vol. 32, no. 3, pp. 353-369.

Staley, Kent W. (2004): "Robust Evidence and Secure Evidence Claims", *Philosophy of Science,* vol. 71, pp. 467-488.

Sugden, Robert (2000): "Credible Worlds: The Status of Theoretical Models in Economics", *Journal of Economic Methodology,* vol. 7, no. 1, pp. 169-201.

Tabuchi, Takatoshi & Thisse, Jacques-Francois (2002): "Agglomeration and Trade Revisited", *International Economic Review,* vol. 43, no. 2, pp. 409-435.

Teller, Paul (2001): "Twilight of the Perfect Model Model", *Erkenntnis,* vol. 55, no. 3, pp. 393-415.

Weisberg, Michael (2006a): "Forty Years of 'the Strategy': Levins on Model Building and Idealization", *Biology and Philosophy,* vol. 21, pp. 623-645.

Weisberg, Michael (2006b): "Robustness Analysis", *Philosophy of Science,* vol. 73, no. 5, pp. 730-742.

Wimsatt, William C. (1987): "False Models as Means to Truer Theories", *in Neutral Models in Biology*, eds. M.H. Nitecki & A. Hoffman, Oxford University Press, .

Wimsatt, William C. (1981): "Robustness, Reliability and Overdetermination", *in Scientific inquiry and the social sciences*, eds. M.B. Brewer & B.E. Collins, Jossey-Bass, San Francisco, pp. 124-163.

Wimsatt, William C. (1980): "Randomness and Perceived Randomness in Evolutionary Biology", *Synthese,* vol. 43, pp. 287-329.

Woodward, Jim (2006): "Some Varieties of Robustness", *Journal of Economic Methodology,* vol. 13, no. 2, pp. 219-240.