Derivational Robustness, Credible Substitute Systems and Mathematical Economic Models: The Case of Stability Analysis in Walrasian General Equilibrium Theory*

D. Wade Hands
Department of Economics
University of Puget Sound
Tacoma, WA, 98416 USA
hands@pugetsound.edu
00-1-253-879-3592
Fax 00-1-253-879-3500

November 2015
Version 2.1

Abstract: This paper supports the literature which argues that derivational robustness can have epistemic import in highly idealized economic models. The defense is based on a particular example from mathematical economic theory, the dynamic Walrasian general equilibrium model. It is argued that derivational robustness first increased and later decreased the credibility of the Walrasian model. The example demonstrates that derivational robustness correctly describes the practices of a particular group of influential economic theorists and provides support for the arguments of philosophers who have offered a general epistemic justification of such practices.

*Helpful comments on earlier drafts were received from Aki Lehtinen, Caterina Marchionni, and two anonymous referees.

> *[The] triumphs of economic model building can evoke suspicion. Do they tell us anything about the world of genuine people, work, production, and commerce? The assumptions of the models often do not seem even remotely accurate as descriptions of an actual economy, … Perhaps economic models, despite their apparent power, turn out to be mere exercises in mathematics … Can models with unrealistic assumptions … be of any use in understanding the world? (Gibbard and Varian, 1978, pp. 664-65)*

0. Introduction

Modern economics has always offered a challenge to those concerned with the epistemic justification of economic science. On one hand, economics seems to be a highly successful special science. Economic theory is extremely systematic, highly mathematical, and backed up by sophisticated statistical techniques employing empirical data from a wide variety of reliable sources. It is also clear that economic theory is respected by those outside of the economics profession since it plays a key role in a wide range of individual and institutional decision-making. And yet, much of economic theorizing, seems to be unnecessarily abstract, insufficiently testable and/or tested, based on extremely unrealistic assumptions, and generally epistemically questionable. The philosophy of a science like this is indeed a challenge.

Although philosophers and economists have been trying to reconcile aspects of these tensions since the nineteenth century, there have been some recent changes in the way such issues are approached. One change has been a naturalistic turn that emphasizes fidelity to the actual practices of economists while at the same time remaining concerned with justification. A second major change, and one that has also occurred in the philosophy of biology, has been decreased emphasis on theories and a corresponding increased emphasis on models. Recent research has made it clear that economic models come in a wide variety of different forms, play a number of roles that are quite different from the roles they play in the physical sciences, and are generally more autonomous with respect to both theory and evidence than previously recognized (Boumans 2005; Knuuttila 2005; Morgan 2001, 2012; Morgan and Knuuttila 2012).

One of the major themes in this recent discussion of economic models has been the question of representation and whether models need to adequately represent real world phenomena in order to give us scientific knowledge. The traditional representationalist view is that if a model is to provide scientific knowledge and acceptable explanations it must, at least to some degree, adequately represent the objects, features, mechanisms, and causal forces that exist, and are at work in, the targeted portion of the real world. Models can suppress various disturbing or inessential factors, but they must adequately represent the fundamental causal forces at work in the target domain.

The most common representationalist view within the recent literature on economic models is the isolationist (or idealization-based) account.[1] According to this view, economic models involve isolating the essential causal forces at work within the relevant target system by screening off various adventitious or disturbing factors. If one thinks of isolation as a process, it starts from a base (the target) and produces a product (a model, or a portion of a model) by isolating various aspects of the target. In the laboratory sciences this isolation or screening-off role is done materially by the laboratory environment, but, it is argued, theoretical models are systematic thought experiments that play a similar isolating role. Since models are based on the relevant portion of the real world – they just neglect certain aspects and distort others – they can serve as representative systems and do real (and realist) scientific work.[2]

Although the isolationist account seems to justify the epistemic significance of, and underwrite the explanations provided by, some idealized models, there are many cases where it is not applicable because economic models fail to represent real world targets in the way the isolationist account requires. Many of the assumptions in economic models not only seal off or distort various disturbing factors, they employ mathematical structures that seem to be totally fictitious and do not connect up in any systematic way with any real world economic targets; they have been called *substitute*, as opposed to *surrogate*, systems.[3] In addition, most of the results produced by manipulation of such models either do not specify any target at all, or specify one that is only a *possible* world that is not part of any real economic world (Grüne-Yanoff, 2013, p. 850). This has led to various efforts to develop nonrepresentational accounts – *fictionalist* positions – that attempt to justify (at least some of) the modeling practices of economists on

---

[1] There is not complete agreement on how exactly the terms "idealization" and "isolation" should be used in these discussions. See Morgan (2012, pp. 137, 158) and Morgan and Knuuttila (2012, p. 51) on this issue.

[2] Influential defenders of the isolationist account in the recent literature on economic models include Nancy Cartwright (e.g. 1991, 2007, 2009) and Uskali Mäki (e.g. 1994, 2005, 2009a, 2009b, 2013).

[3] A *substitute system* is Mäki's term for a model that is "a freely floating subject of inquiry, unconstrained by any real concern as to how it might be connected to real world facts" and they are "*strong failures* of representation" (Mäki, 2009b, p. 36). For Mäki the alternative to a substitute system is a *surrogate* system that adequately represents (or at least is intended to represent) the relevant target, but many would argue there is a third alternative – I will use the term *credible-substitute models* – that do tell us something about certain aspects of the economic world. We learn from them, make inferences with them, and assess credibility-increasing and credibility-decreasing moves within them, while they still tell us less about real economies than fully representational models.

grounds other than representational veracity. These accounts replace representational surrogate models with fictional models that create parallel (not representational) worlds and yet still convey useful knowledge about various aspects of the economic world. The economist Robert Sugden calls such models *credible*. As he explains:

> We recognize the significance of the similarity between … model markets and real markets, by accepting that the model world *could be* real – that it describes a state of affairs that is *credible*, given what we know (or think we know) about the general laws governing events in the real world. On this view, the model is not so much an abstraction from reality as a parallel reality. The model world is not constructed by starting with the real world and stripping out complicating factors: although the model world is simpler than the real world, the one is not a *simplification* of the other. (Sugden, 2000, p. 25).

As a result of these differences, there is an on-going debate about the adequacy of these competing views:[4] *isolationist* versus *fictionalist*[5] accounts of theoretical modeling in economics.[6]

So if many economic models fail as adequate representations, then what tools are available to help us sort out models that tell us something about the economic world, and perhaps can be relied upon for successful interventions in that world, from those that are purely substitute systems? The suggestion that will be examined here is an idea that goes back to at least Gibbard and Varian in 1978 (p. 774): the notion that *robustness* can, at least in part, serve this function. A robust model is one where the same result can be derived from a wide range of different auxiliary assumptions. The general idea is that if model X is a combination of a core model C and a set of auxiliary assumptions $A_1, \dots, A_n$ and X implies the

---

[4] The word "competing" needs to be clarified here. Most who support some version of a fictionalist account would agree that certain economic models do satisfy the requirements of the isolationist account. The problem is that many do not, and these models are often the most influential work economists produce. The problem is not that the isolationist view is incorrect, it is simply that it is not applicable to important theoretical modeling in economics, some of which, the defenders of fictionalist views assert, can also be justified. In other words, isolation may be sufficient for the success of models, but it is not necessary (Grüne-Yanoff, 2011; Knuuttila 2009, 2011). Perhaps the best way to think about the issue is to draw an analogy with a much earlier debate in the philosophy of science. Ian Hacking once said that Popper took the positivist dichotomy of science vs. metaphysics/muck and converted it into the three-way distinction of science vs. metaphysics vs. muck (Hacking, 1979, pp. 384-385). Similarly, we could replace the isolationist dichotomy of surrogate vs. substitute/muck with an alternative three-way distinction of surrogate vs. credible substitute vs. muck.

[5] The *fictionalist* view is sometimes called *constructionist*, but I will employ only the former term to avoid confusion with constructionist views within the sociology of science.

[6] Contributions to this literature include: Cartwright 2009; de Donato Rodrïguez and Zamora Bonilla 2009; Grüne-Yanoff 2009a, 2009b, 2011, 2013; Knuuttila 2005, 2009, 2011; Kuorikoski and Lehtinen 2009; Mäki 2009b, 2013; Morgan and Knuuttila 2012; Sugden 2000, 2009; and Weisberg 2007, 2013.

result R, so $X = (C, A_1, A_2, \ldots, A_n) \Rightarrow R$, then X is robust if it can be changed by modifying one of its auxiliary assumptions to say $X' = (C, A'_1, A_2, \ldots, A_n)$, or $X'' = (C, A''_1, A_2, \ldots, A_n)$, etc. and still continue to imply R.[7] Robustness is thus a kind of stability result and the opposite of a robust model is one whose results are sensitive to small variations in the assumptions and are thus unstable or fragile.[8]

There are many different kinds of robustness analysis and the one that is relevant to the mathematical models of interest here is *derivational robustness*. Broadly, derivational robustness is robustness of a theoretical result with respect to particular variations in the model's auxiliary assumptions. If we think of a model with clear empirical implications, then these changes are what testing the model in different laboratory environments, different sample populations, different measuring devices, etc. accomplishes. The argument is that such robustness testing can play a similar epistemic role in the thought experiment world of mathematical models that have deductive implications that are not observational in any straightforward way. As Jaakko Kuorikoski and Aki Lehtinen describe it:

> Derivational robustness analysis is the procedure for testing whether a modelling result is a consequence of the substantive assumptions or an artifact of the errors and biases introduced by the auxiliary assumptions. It is carried out by deriving a result from multiple models that share the same substantive assumptions but have different auxiliary assumptions. (Kuorikoski and Lehtinen, 2009, p. 126)

With this introduction it is now possible to state the purpose of this paper. The paper examines the credibility-enhancing capabilities of derivational robustness in mathematical economic models. The argument will be made by means of an extended economic example: the stability literature in dynamic Walrasian general equilibrium theory from the middle of the twentieth century. It will be demonstrated that derivational robustness was a main reason why economists initially thought the stability literature was positive – why it increased confidence in the Walrasian general equilibrium model – but also why later, after proofs demonstrating that additional robust results would be impossible, it was interpreted quite negatively.

---

7 This is a modified version of the characterization of robustness in Odenbaugh and Alexandrova, (2011, p. 764).

8 The recent literature supporting various types of robustness analysis in economics includes: Guala and Salanti 2003; Kuorikoski and Lehtinen 2009; Kuorikoski, Lehtinen, and Marchionni 2010, 2012; Lehtinen and Kuorikoski 2007; Lehtinen and Marchionni 2011; Weisberg 2013; and Woodward 2006. For critical remarks on robustness analysis see Cartwright (1991, 2007, 2009), Odenbaugh and Alexandrova (2011), and Reiss (2012). It should be noted that robustness analysis is also an important topic in the philosophy of biology (particularly in systems biology and ecology) and many of the papers discussing robustness in economics also consider biology. I will only be concerned with economics.

1. Derivational Robustness, and Increased Credibility in Non-Representational Models in Economics

It has been argued that certain types of robustness analysis can provide empirical support for models that have clear observable implications. However, it is important to remember that this is not the case for the type of models of interest here. The implications of the mathematical models considered here are simply *results*.[9] They are various propositions involving equilibrium conditions, sign restrictions, or as in the example discussed below, the stability property that the dynamic price path p(t) converges to the equilibrium as $t \to \infty$. For the remainder of this paper robustness will refer exclusively to *derivational robustness* in this type of mathematical economic model.

So how exactly can such robustness analysis increase the credibility of economic models? The most sustained defense has been given by Jaakko Kuorikoski, Aki Lehtinen, and Caterina Marchionni (2010, 2012). Their goal is to "defend the practice of derivational robustness analysis against the accusation of epistemic sterility" (ibid., p. 549), and they offer a number of interrelated arguments for how robustness can have epistemic import.

Their general argument is that it seems highly unlikely that the result would continue to hold for such a wide range of auxiliary assumptions if the basic causal mechanism identified in the model were not responsible for the result.

> … 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 conclusions if the conclusion did not correspond to something real. (Kuorikoski, Lehtinen, and Marchionni, 2010, p. 544)

The more specific aspect of their defense focuses on the ability of robustness to provide information about the impact of particular auxiliary assumptions.

> Our claim is that independence of a modelling *result* with respect to particular modelling assumptions may nonetheless carry epistemic weight by providing evidence that the result is not an artefact of particular idealizing assumptions. In particular, we argue that

[9] "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" (Kuorikoski, Lehtinen, and Marchionni, 2010, p. 545).

> 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 … (ibid., pp. 542-543).

Since different types of assumptions play very different roles in economic models, it is useful to review the three types of assumptions that Kuorikoski, Lehtinen, and Marchionni discuss: substantial, Galilean, and tractability (ibid., p. 547). Substantial assumptions identify the core causal mechanisms that are considered to be responsible for the results in question. Galilean assumptions serve isolationist functions "by idealizing away the influence of the confounding factors" (ibid., p. 547). Finally, tractability assumptions are assumptions known to be false, but added to increase the mathematical tractability of the model; examples from economics include the representation of bundles of commodities by vectors of real numbers and the representation of the technological relation between the inputs and the output of a good by a differentiable production function. For Kuorikoski, Lehtinen, and Marchionni robustness involves all three kinds of assumptions, but tractability assumptions are the most epistemically problematic and the most relevant to derivational robustness analysis (2010, p. 548). This applies to both the case where the analysis has a positive outcome as well as where it fails.

To see how positive robustness results work recall the core and auxiliary symbolism introduced above, where model X was a combination of a core model C and a set of auxiliary assumptions $A_1, \ldots, A_n$ which implied the result R, but now suppose the first auxiliary assumption is a tractability assumption. Changing the symbol to T for tractability, and assuming there are m different tractability assumptions that could be used, we have m different models $X^i = (C, T_i, A_2, \ldots, A_n)$ for all i = 1, …, m. Let us also assume that each of these different tractability assumptions are independent in the sense that $T_i \nRightarrow T_j$ and $T_j \nRightarrow T_i$ for all $i \neq j$. If the result is derivationally robust with respect to these tractability assumptions then $X^i \Rightarrow R$ for all i. In the case where the robustness test is successful, knowing that the result R is not fragile with respect to this range of tractability assumptions gives us more confidence that R is a result of the other aspects of the model $(C, A_2, \ldots, A_n)$. While this seems reasonable, it is also clear that robustness testing may not be this simple. Auxiliary assumptions, even pure tractability assumptions, often have various kinds of interdependencies. So for instance instead of, say, T_i being a single assumption it could be a cluster of different tractability assumptions. Perhaps $T_i = (T_i^1, T_i^2, \ldots T_i^{k(i)})$ for all i = 1, … , m and some k(i) for each i. The number of possible tractability assumptions in the cluster, k(i), is written as a function of i, since there is no reason that each cluster would have the same number of assumptions. As the stability example in the next section demonstrates, this case seems more likely to fit the actual modeling practice in economics, but it is messier. Obviously this case reduces to the previous case when all of the assumptions in each cluster are independent of all of the assumptions in all of the other clusters, but that seems unlikely when fairly general mathematical assumptions are involved such as continuity or differentiability. Still, even if this strong case does not hold, it may be that as long as some of the assumptions in each cluster are independent of the assumptions in all of the other clusters, or if we have additional information about the overlapping assumptions, the derivational robustness analysis operates in the same way as the simple case where they are all independent. Of course a direct empirical confirmation would be better, but in the models that concern us here, this is almost never an option; if direct empirical tests were available, robustness analysis of this type would be much less important (Kuorikoski, Lehtinen, and Marchionni, 2010, pp.552-53; 2012, p. 895).

But now consider the case when the robustness analysis fails, where the result is sensitive to variation in the tractability assumptions. Suppose $X = (C, A_1, A_2, \ldots, A_n) \Rightarrow R$ but $X' = (C, A_1', A_2, \ldots, A_n) \nRightarrow R$. Here we have a (non-empirical) version of the Duhem-Quine problem. As always in such situations a *decision* is required; in this case there seem to be two main ways to respond to the robustness failure. First, suppose there is good reason to believe that A_1 really is just a tractability assumption. In that case, it means that "a tractability assumption drives the result" (ibid., p. 553) not other, more substantive, portions of the model. This means a decision needs to be made about how the model should be substantively revised (or perhaps even rejected). A second option is to consider the possibility that A_1 is not simply a tractability assumption; that the model depends on it in a more substantive way. In either case the evaluation "requires looking at and attributing different degrees of credibility or reliability to *parts* of models" (ibid., p. 551).

In the next two sections I will apply this analysis of derivational robustness to an important set of economic models: the stability of equilibrium in a Walrasian general equilibrium model. I will argue that the aspects of robustness emphasized by Kuorikoski, Lehtinen, and Marchionni were important in the stability literature – and that economic theorists responded to the success and failure of robustness in the ways their analysis suggests – but I will also argue that there is one aspect of this example that goes beyond their discussion.

2. The Example of the Stability Literature in Walrasian General Equilibrium Theory (1945-1975)

This section will lay out the basic structure of the Walrasian general equilibrium model that will be used to discuss the relationship between robustness and stability theory. Given the naturalistic spirit of the recent literature on economic models, almost every author employs specific examples from economic theory to help make their case; one much-discussed example is Schelling (1978). Such examples serve their desired purpose, but I believe an example from Walrasian general equilibrium theory may be more effective for the issues under consideration here. For one thing, general equilibrium theory was, and to some extent still is, the theoretical core of rigorous economic theory – including being the foundation of graduate education during the last half of the twentieth century – and while Schelling's model is (justly) respected, it has always been considered a special topic and was never anything like the core analytical framework of economic theory. Secondly, models such as Schelling's often start from a general feature of the real world (although not necessarily a specific real world event) that prompts an explanation (Aydinonat 2007, 2014; Elliott-Graves and Weisberg 2014; Weisberg 2013) and this concretizes such models in a way that was never the case for abstract general equilibrium theory. Finally, I will show how robustness played a role both in the period where the stability results were considered successful and the later period where they were considered a failure, thus using one example to examine both aspects of derivational robustness analysis.

The name Walrasian comes from Léon Walras who provided the most systematic early formulations of general equilibrium theory in the 1870s (Walras 1954), and although the theory originally presented by Walras had much in common with the version discussed here, there are also differences. I will explain the features of modern Walrasian general equilibrium models that are key to understanding the role that robustness played in the development of such models, but before doing so it is useful to discuss this general category of economic models. Instead of starting immediately with the Walrasian model, I will start with the more general idea of a *perfectly competitive economy*. The perfectly competitive economy is a model, but it is a very general conceptual model, not a tightly structured mathematical model like the Walrasian model. The Walrasian model is one specific instantiation of the more general concept of a perfectly competitive economy. Lehtinen and Marchionni (2011, p. 4) use the term "platform" for a quite general model and "sub-model" for a specific instantiation of that general model. In their terminology, the perfectly competitive economy would be the platform and the Walrasian general equilibrium model a particular sub-model.

A perfectly competitive economy is an economic world where every firm (seller) and every consumer (buyer) takes prices as parameters, not choice variables. As an example of a real world firm that approximates perfect competition, think of a wheat farmer. There are millions of wheat farmers, so many in fact that no individual producer has any control over the price of the wheat they sell. They grow and harvest it and sell it at *the market price*. If they think this price is too low they can refuse to sell, but the price will not change if they do so because there are millions of other wheat growers in the world producing essentially the same product. The perfectly competitive firm is thus a *price taker* (not price maker), or thinking more mathematically, price is a parameter (not a choice variable). Consumers act the same way – they take prices as parameters – and simply decide to buy or not at the market price. If both buyers and sellers take prices as parameters then all markets are perfectly competitive and an economy where all markets are perfectly competitive is a perfectly competitive economy.

Notice that an actual perfectly competitive economy is an economy that has never existed nor ever will exist; it is a hypothetical economy with no monopolists, no oligopolies, no production by the government or non-profits, no brand names, free exit and entry into every industry, and a host of other features. And yet it is a *possible world*. It does not violate any laws of logic or nature, and a few sectors of real market economies even approximate it (like agriculture). Perhaps a useful way to think about it is in the way that Mary Morgan characterizes Max Weber's ideal types: "generalizations constructed from the 'facts of experience' yet in the process, creating abstract concepts that he described as 'pure fictions'" (Morgan, 2012, p. 142). Perfectly competitive economies function "neither as theories or empirical descriptions, but as independent instruments or tools that enable the social scientist to support enquires into both domains" (ibid., p. 144).[10]

There are many different ways of modeling a perfectly competitive economy, each sharing the core features of price-taking behavior and equilibrium, but each having its own particular features.[11] The competitive equilibrium model of interest here is the Walrasian general equilibrium model developed during the 1950s – often called the Arrow-Debreu model because of the key existence result Arrow and Debreu (1954) – and became the core microeconomic theory during the 1960s. It has many variations – including pure exchange and production versions – but such details are not relevant to the argument here. The canonical

[10] The assumptions of a perfectly competitive economy also demonstrate how difficult it is to classify the various types of assumptions used in economic models (i.e. how a single assumption can play multiple roles). The key assumption of price-taking behavior is a substantive assumption about the behavior of the agents in the model (and is approximated by some real world agents), yet it is also a very effective tractability assumption since taking prices as parameters greatly simplifies the application of differential calculus to such models.

[11] Attention was given to perfect competition in the work on idealization in economics published in the 1980s and 1990s (e.g. Hamminga 1983 and various papers in Hamminga and De Marchi 1994), but it has received much less attention in the recent literature. Mäki (2001) discusses perfect competition, but from the viewpoint of three economists who criticize it.

text is Debreu (1959), although Arrow and Hahn (1971) is a better reference on the stability literature.[12]

The general version of the model has n goods, where n could be any finite number. In the production case the supply of each good is given by profit-maximizing perfectly competitive firms and the demand for each good is given by budget-constrained utility-maximizing consumers. The objective function of each firm (profit) is based on the firm's technology and the prices they face, while the objective function of each consumer (utility) is based on the consumer's preferences and constrained by their available income. The primitives are technology, tastes (preferences), and income/endowments. Optimizing over these primitives generates the supply and demand for each good, and the equilibrium of supply and demand in each market determines the equilibrium of the economy. The n-good Walrasian model is far too conceptual and abstract for direct empirical testing of its assumptions, but the general idea that there are two sides to a competitive market (producers/sellers and consumers/buyers), that the actions of consumers are based on what they prefer and can afford, and the actions of firms are based on profit, seems to be consistent with generally accepted stylized facts about market economies. This is not empirical content in the traditional sense, but it is a (weak) empirical link to the real economy. The model represents, but it does not directly represent any particular part of the economic world; as noted above, it represents conceptual content which is, in some indirect sense, constructed from the "facts of experience."

If we let $p = (p_1, p_2, \ldots p_n)$ be the prices of the n goods, the function $S_i(p)$ the supply of good i, and $D_i(p)$ the demand for good i, then the equilibrium price vector $p^* = (p^*_1, p^*_2, \ldots, p^*_n)$ will solve the system of equations $D_i(p^*) - S_i(p^*) = 0$ for all i.[13] Underlying these supply and demand functions are of course the tastes, technology, and endowments, but for our purposes it is not necessary to write out the optimization problems of either type of agent. If we define the excess demand (Z) for good i by $Z_i(p) = D_i(p) - S_i(p)$, then equilibrium is given by $Z_i(p^*) = 0$ for all i. Under these standard assumptions the Walrasian model has two important implications: Walras' Law (W) and zero degree homogeneity of excess demand (H):

$$\sum_i p_i Z_i(p) = 0 \qquad \text{(W)}$$

$$Z_i(\lambda p) = Z_i(p) \text{ for all p and } \lambda > 0. \qquad \text{(H)}$$

This is essentially the Walrasian general equilibrium model. There were many expanded versions of this basic model that added special assumptions – from

[12] McKenzie (2002) provides a fairly recent survey of this type of general equilibrium theory, Quirk and Saposnik (1968) is a text from the period that emphasized stability, and Hahn (1982) provides a useful survey. See Ingrao and Israel (1990) for historical discussion.

[13] For our purposes we can assume $p^* > 0$.

money, to public goods, to capital goods, etc. – but this core structure was preserved in essentially all of these special cases.

The basic Walrasian model is an equilibrium model, but it contains no mechanism that might explain how prices get to equilibrium. It would seem to be an improvement in the model to include a dynamic price adjustment – to show how (or how possibly) prices might adjust, and hopefully converge to, the equilibrium price vector p* over time. There are several reasons why such a "price adjustment mechanism" might increase the credibility of the Walrasian model, and thus the confidence in the Walrasian model as the best characterization of a perfectly competitive economy: I will note just three.

i) *Comparative Statics*: Most of the "analysis" that economists do with mathematical models involves what economists call "comparative statics." The basic idea is that one starts out with an equilibrium model characterized by a system of equations (say $H_i(x^*, \beta) = 0$ for all i) involving both variables (say x_is) and parameters (say β_is). When one of the parameters changes ($\Delta\beta_i$), the equilibrium values of (at least some of) the variables will change (Δx_i^*). If the model has sufficient mathematical structure it will be possible to determine (at least the sign of) the change in the equilibrium values. The problem is that one needs to know that after the change in a parameter the model will reach a new equilibrium that *can* be compared to the original equilibrium. Samuelson originally called this relationship between stability and comparative statics the *Correspondence Principle* (Samuelson, 1947, pp. 258, 284, 350). Thus adding a price adjustment mechanism to the static Walrasian model would help rationalize the comparative statics exercises economists perform with such models.

ii) *Logic*: There seems to be a logical problem with the Walrasian model; it is not technically a contradiction among the various assumptions, but the model does seem to have a gap. The model assumes that both firms and consumers "take prices as parameters" so prices are not determined by the agents in the model. But these are the only agents in the model. So where do prices come from? If everyone "takes" prices, then who "gives" them? The professional folklore of economists would suggest the "invisible hand" of the market, but there is no such mechanism in the basic Walrasian model (Arrow 1959). Adding a price adjustment mechanism adds an *institution* to the model that fills in this logical gap.

iii) *Stylized Facts*: As noted above, while abstract mathematical models like the Walrasian model are not directly empirically testable, the credibility of the model is improved when it is consistent with certain generally accepted stylized facts. One such "fact" is that if, at any price, the quantity that buyers want to buy is greater than the quantity that sellers want to sell, then the price increases; and that if, at any price, the quantity that buyers want to buy is less than the quantity that sellers want to sell, then the price decreases. There are situations where this is not the case of course, but they are rare. In general this is a stylized fact of markets that holds over different times, places, and institutional configurations (e.g. both humans and software trading financial assets). It would add to the credibility of the Walrasian model if this stylized fact could be accommodated.

So how was the Walrasian model modified to accommodate price adjustment? Walras himself offered an adjustment mechanism he called *tâtonnement* and thus introduced the term that became standard in the later literature, but the *tâtonnement* mechanism that will be discussed here came from Samuelson's papers on *true dynamic stability* in 1941 and 1942. Here Samuelson characterized the price adjustment process as a system of ordinary differential equations where prices changed through time (t) in the direction of excess demand. Let us call this *tâtonnement* process (T):

$$dp_i / dt = Z_i(p) \text{ for all i.} \qquad \text{(T)}$$

Notice that given the earlier definition of Z, this mechanism captures the stylized facts of price movements in competitive markets – positive excess demand for a good increases the price and negative excess demand (excess supply) decreases the price – and it does so in a mathematical framework that is fairly easy to work with.[14]

So how does (T) relate to price changes in the real world? Of course the real world contains firms with pricing power who decide prices, but let us disregard that and ask the question of how (T) relates price changes in competitive markets. On first gloss it seems to represent these changes reasonably well – it is consistent with the stylized facts – but on closer examination there are serious concerns. For one thing, to employ (T) and be consistent with the foundational assumptions of the Walrasian model, economists assumed there would be no trading outside of equilibrium. If $Z_i \neq 0$ then either there will be buyers who cannot buy or sellers who cannot sell. If trade took place at such disequilibrium prices the underlying optimization problems would need to be recomputed and the demand and supply functions themselves would change. Explaining this in more detail:

> … in terms of our equations, we defined Z(p*) = 0, and then described the dynamic process by dp(t)/dt = Z[p(t)]. In other words, we used the same function Z to denote the equilibrium relation and the dynamic process. If we allow intermediate purchases and actual transactions in the process, then this excess

[14] Other specifications were used in the literature such as $dp_i/dt = k_iZ_i(p)$ with $k_i > 0$ and $dp_i/dt = H_i[Z_i(p)]$ with $H_i' > 0$, but these differences can be neglected here. These, like (T), are systems of ordinary differential equations where prices change in the direction of excess demand.

> demand function Z will change from the time to time as the trader's income or purchasing power varies. Hence the price vector which prevails when the market is finally cleared depends on the time path of the process and will, therefore, not generally be the same for any two processes. Thus, the process does not describe *at all* how the economy actually reaches an equilibrium price vector p*, the very problem with which Walras was concerned. (Takayama, 1974, pp. 341-42, emphasis added)[15]

Perhaps the best way to think about (T) is as an auctioneer's rule for how to change prices:

> We might imagine, to give some flesh to the abstraction we propose to investigate, the existence of a super-auctioneer who calls a given set of prices p and receives transaction offers from the agents in the economy. If these do not match, he calls another set of prices, … but no transactions are allowed to take place. This process either comes to some end or continues indefinitely. (Arrow and Hahn, 1971, p. 264)

The empirical relevance of such an auctioneer's rule is unclear. It undoubtedly captures the rule followed by some real auctioneers, but the vast majority of markets do not have auctioneers and real markets involve exchange at non-equilibrium prices. But (T) gets even more problematic as we examine it further. Under (T) the equilibrium p* is *stable* (discussed in detail below) if the price path p(t) generated by (T) converges to the equilibrium: in other words, that limit p(t) → p* as t→ ∞. Since there is no trading until equilibrium and equilibrium is only reached when t→∞ then it seems that trading would never occur – which is not only unrealistic, it also suggests the (T) is *not even in principle a mechanism that could possibly produce competitive prices* in finite time. As Arrow and Hahn explain:

> … we insisted that no trade take place out of equilibrium. This restriction, strictly interpreted, is not only obviously unrealistic, but also seems to carry the logical implication that trade never takes place. If the auctioneer's rule is not stable, trading is not permitted *a fortiori*, while if it is, trading will be permitted only "in the limit" (i.e. as t approaches infinity), for it is only in the limit that equilibrium and "called" prices coincide. (Arrow and Hahn, 1971, p. 324).[16]

---

[15] Takayama's symbolism was changed to be consistent with the symbolism used in this paper.

[16] I note in passing that these concerns did motivate economists to try to develop models without some of the undesirable features of (T) – in particular to allow for disequilibrium trading – but the research project was never very successful: see chapter thirteen of Arrow and Hahn (1971) for a survey of the results at that time.

Given all of these concerns it is clear that the dynamic Walrasian general equilibrium model – the basic model with (T) appended – is *not a surrogate representational model*. There are too many abstractions, too many unrealistic assumptions that are fictions rather than isolations, and the model's target is a conceptual model (the perfectly competitive economy) that is only indirectly linked to the real economic world. Perhaps most important of all, not only is there no reason to believe that (T) captures the underlying causal mechanism of competitive price determination, it is pretty clear the economists who formulated it never *intended* it to do so nor *believed* that it did so. Economists of course did (and still do) not know exactly what causes prices to change, but the general intuition of economists (then and now) is that prices change as a result of the uncoordinated actions of individual agents (firms and consumers)[17] not a super-auctioneer or any other centralized mechanism.

But it is also clear that it is not a *pure substitute model either*. First of all, (T) is consistent with the relevant stylized facts of competitive price adjustment. In addition, the general idea of a perfectly competitive economy as well as the specifics about the role of supply and demand in competitive markets, has guided how economists thought about real market economies since the end of the eighteen century and the framework cut across ideologies as well as time (from Adam Smith to Karl Marx to Léon Walras to Paul Samuelson). The Walrasian theoretical framework including (T) – particularly in Lange (1944) and Patinkin (1965) – also served as the microeconomic backdrop for the version of Keynesian macroeconomics that was dominant during the middle of the twentieth century: a model that had (for a while) a successful record of real world policy *interventions*. (T) was appended to the basic Walrasian model and that model clearly involved causal processes that economists believed to be at work in real economies – particularly profit and utility maximization – and these were mechanisms that economists had tried to test with empirical data for many decades (with mixed results). Economists also tried to test the main implications of the basic Walrasian model – (W) and (H) – on specific market data, again with mixed results, but serious attempts to test these implications suggest that these conditions were considered to be something that was not merely mathematical but rather something that could potentially represent a part of the real economic world.[18] Finally, while perfectly competitive behavior is rare in firms, it is a good approximation to firm behavior in a few sectors, and for consumers it is the norm

[17] This belief can be thought of as what Mäki calls the ontological *way the world works* (www) constraint on scientific theorizing. Various scientific communities (and research programs within those communities) have commitments to particular "causal processes that constitute the ways the world works" (Mäki, 2001, p. 371) and for most economists that commitment was (and is) to the rational actions of individual agents, not central planners or mechanisms that represent the behavior of a central planner, as the ultimate cause of market phenomena.

[18] See Hands and Mirowski (1998) and the references therein for a discussion of these various testing efforts.

(as consumers we normally just take, rather than make, the prices of the goods we purchase).

So the bottom line is that the dynamic Walrasian model is *neither surrogate nor substitute*. It is a *credible-substitute* system conditioned by (some) of the facts about the world and also by (some) of the causal mechanisms that economists believe to be at work in the economy, and yet it does not represent – in fact, approximately, or with respect to the intentions of the relevant economists – any real economic world.

I have now introduced three interconnected models:

i) The general conceptual model of a perfectly competitive economy,
ii) The Walrasian general equilibrium model (in basic, or static, form),
iii) The Dynamic Walrasian general equilibrium model [model ii) with some version of (T) appended].

So what was the relationship among these three models? How did the economists working with them think about the relationship and how should one think about the relationship in order to understand how robustness worked to increase the credibility of the Walrasian model?

I argue that the first two of three models are *nested*; the concept of a perfectly competitive economy is the most general and the Walrasian model is one particular way to model a perfectly competitive economy, while the dynamic Walrasian model is an *extension* of the basic Walrasian model that adds the dynamic component (T)[19]. I will avoid the controversial issue of the epistemic value of the general concept of a perfectly competitive economy. I will simply assume, reasonably I think, that economists during the period believed the concept of a perfectly competitive economy to have epistemic import: it told them something about real markets (but I will not attempt to specify exactly what that was). So too for the Walrasian model. I will assume there is some initial positive level of commitment to the credibility and epistemic value of the Walrasian model as the best model for analyzing perfectly competitive economies. I will just assume these starting levels of credibility and epistemic value and focus on the question of how the various results derived from (T) *increased or decreased* that initial level. The question is not why economists had (or whether they should have had) confidence in the Walrasian model, but rather why a series of robustness results first increased and then decreased that confidence. I realize this approach – focusing on directional changes rather than defending particular initial values – is a rather modest approach to either

[19] One might think about this as an example of the *hierarchy of models* (Suppes 1962, 1967), but I will not pursue this idea here.

description or justification, but it seems to be the only way to keep the discussion to manageable proportions.[20]

In summary, I will argue that derivational robustness analysis on the dynamic Walrasian model increased the credibility of the basic Walrasian model, which in turn increased confidence that it was possible to learn about the real economy from perfectly competitive models. The next section will examine the way that economists used robustness analysis in the dynamic Walrasian model through three different periods of research.

3. Models That Live by Robustness, Can Die by Robustness

Before working through the research of the three periods it is useful to say a bit more about both the mathematics and the various types of assumptions involved in the dynamic Walrasian general equilibrium model.

The stability concept used in the Walrasian literature is *asymptotic* stability. For any initial (non-equilibrium) price vector $p(0) = [p_1(0), p_2(0), \ldots, p_n(0)]$ the stability question is whether the price path $p(t) = [p_1(t), p_2(t), \ldots, p_n(t)]$ generated by (T) will converge to $p^* = [p_1^*, p_2^*, \ldots, p_n^*]$. If $\lim p(t) = p^*$ as $t \rightarrow \infty$ then the system is asymptotically stable in the sense that the prices will converge to p* (i.e. p* is a sink or attractor). But such convergence could be *local* or *global*. *Local stability* says that the price path converges for initial values p(0) within some epsilon neighborhood of p* while *global stability* says that the price path converges for p(0) anywhere in the price domain. Global stability is obviously a much stronger result than local stability.

The technique used to analyze local stability was the technique popularized by Samuelson. Using (H) and taking the Taylor series expansion of the function Z(p) it is possible to linearize the differential equation system (T) into the matrix form:

$$dp/dt = JZ(p^*)[p(t) - p^*], \qquad (T_L)$$

where JZ(p*) is the Jacobian matrix of the excess demand functions evaluated at p* with representative term $Z_{ij} = \partial Z_i(p^*)/\partial p_j$. ($T_L$) is stable if the real parts its characteristic roots are all negative and this is sufficient for the local stability of the original non-linearized system (T). Global stability will be discussed below.

Turning now to the question of different types of assumptions, recall that Kuorikoski, Lehtinen, and Marchionni discussed three types – substantive, Galilean, and mathematical tractability – and emphasized tractability in their

[20] There is also an argument that robustness analysis in general is only relevant to directional changes, but commitment to this view is not necessary for the argument offered here.

discussion of robustness. These are important distinctions, but I would like to discuss an additional useful classification of assumptions: those that have an *economic interpretation*. Although substantive and Galilean assumptions can and often are economically interpretable, the economically interpretable assumptions that are of interest here are tractability assumptions; t*ractability assumptions that are economically interpretable often carry some weight with respect to credibility*. In mathematical economic modeling credibility often comes from *merely having* an economic interpretation. Saying that an assumption has an economic interpretation, or economic meaning, should *not* be confused with saying it is empirically confirmed or empirically accurate. E*conomic interpretability* in this sense only requires two things: not being impossible given generally accepted stylized facts about the economy and/or economic behavior *and* not being inconsistent with the core presuppositions of the relevant theory. Notice that having an economic interpretation involves both an empirical constraint (not impossible given the stylized facts) and a disciplinary constraint (not inconsistent with the core presuppositions of the relevant theory). In the case of the Walrasian model the latter disciplinary condition generally means being consistent with profit- and utility-maximization. For example, assuming that a particular good in a Walrasian model is a so-called Giffen good – a good with an upward sloping demand curve – might be a tractability assumption (a false assumption with useful mathematical properties), but it also has an economic interpretation: although false, it is empirically possible and it is not inconsistent with the core assumptions of the Walrasian model. On the other hand, assuming that *all* goods in a Walrasian model are Giffen might also be a useful tractability assumption, but it does not have an economic interpretation; it violates the underlying assumptions of utility maximization.[21] One good example of an economically interpretable tractability assumption is the *Gross Substitute* (GS) condition discussed below. It restricts the signs of the cross-partial derivatives of excess demand functions to be strictly positive so $Z_{ij} = \partial Z_i(p^*)/\partial p_j > 0$ for all $i \neq j$. This is a mathematical restriction that serves as a tractability assumption – it is empirically false (not all goods are gross substitutes based on the empirical analysis of real demand systems) and it is mathematically quite useful – yet it is *not just* a tractability assumption; it *has* an economic interpretation. GS is not impossible given the available stylized facts and it is not inconsistent with the core optimization assumptions of the Walrasian model. This makes it more credible than a mathematical assumption with the same formal implications that does not have an economic interpretation. This is just one example, but the point is that there are assumptions used in mathematical economic models that serve dual roles – tractability and credibility – and they are introduced precisely

[21] Being a Giffen good means that the good is inferior (consumption decreases with increases in income) and given the budget constraint associated with the standard utility-maximization problem, not all goods can be inferior.

because they serve both roles. Such assumptions need to be taken into consideration.[22]

3a. Robustness-Based Optimism 1945-1960

The publication of Samuelson's two papers on stability in 1941 and 1942 sparked an extensive literature on the local stability of the dynamic Walrasian model, even though there was very little discussion of that model in the two papers. His brief discussion of the Walrasian model emphasized the differences between his own *true dynamic stability* and the stability analysis of John Hicks in *Value and Capital* (1939). Samuelson's papers were very important to the later stability literature, but because they provided mathematical tools that would allow others to engage in robustness analysis of the dynamic Walrasian model, not because of his own analysis.

Virtually all of the results on local stability in the dynamic Walrasian model during the next two decades *were examples of derivational robustness analysis.* The formula for these papers was almost always the same. Start with the dynamic Walrasian model and add an additional assumption – generally one that restricts the matrix Jf(p*) in some way – then demonstrate that the additional assumption implies stability.

One of the early results was stability under the assumption that all goods were gross substitutes (GS). Three authors proved local stability with (GS) in papers published in 1958: Arrow and Hurwicz (1958), Hahn (1958), and Negishi (1958). All three papers were successful robustness results; they all started with the same substantive assumptions (from the static Walrasian model), a version of (T), and GS, and yet each employed slightly different tractability assumptions. But the results were even better since the GS assumption – unlike Samuelson's characteristic root condition or Hicks's conditions – had an economic interpretation. I argue this is how the most successful results worked in this literature. Pure derivational robustness, as long as the core Walrasian assumptions are intact, increased credibility, but if one of the new assumptions was economically interpretable, it was considered to be an even more significant result.

A good example of how extensive the literature on local stability became is Peter Newman's 1959 survey. The paper discussed seven different assumptions, and listed sixteen different theorems derived from these assumptions. There is no reason to review all of these results, but I will note a few examples. Two of the

---

22 Kuorikoski, Lehtinen and Marchionni do recognize that an assumptions that have dual roles – "a single explicitly stated modelling assumption may simultaneously encode a tractability assumption as well as a substantial assumption" (2010, p. 548) – but do not have a separate category for such assumptions or explicitly define the property of having an economic interpretation.

most straightforward results are his theorems three and thirteen; theorem three proves stability when the matrix JZ(p*) has the mathematical property of being quasi-negative definite (QND), and theorem thirteen proves stability under the assumption that JZ(p*) has a quasi-dominant diagonal (QDD). If we think of these conditions as purely mathematical tractability assumptions, then both of these theorems clearly provide successful robustness results for the dynamic Walrasian model. But there is a difference between the two results: QND does not have an obvious economic interpretation while QDD does. QDD says that "own" effects (based on partial derivatives of the excess demand functions) dominate the sum of "cross" effects. As Newman notes: these (QDD) "conditions are obviously restrictive, but equally obviously have clear economic meaning" (ibid., p. 4) and this gives them a slight credibility edge over purely mathematical conditions like QND. But GS has a credibility edge over both of these conditions. GS also has an economic interpretation, but it dominates QDD in two ways; it is considered to be more empirically acceptable[23] (although again this is a relatively weak sense of "empirical") and it is also mathematically stronger since GS ⇒ QDD (Newman's theorem sixteen). There were many local stability results, but the one that came to have the most credibility was the GS assumption. Like all of the other assumptions in the literature it was sufficient for local stability, but it was the assumption with the most acceptable economic interpretation (even though it was never found to hold in empirical demand systems involving many goods) and it also implied some, but not all, of the other assumptions that were sufficient for local stability. By 1968 it could be said that:

> There is no other case of comparative generality in general equilibrium theory concerning which as much is known as the gross substitute case. In a certain sense, it practically exhausts the comparative statics content of general equilibrium analysis …
> (Quirk and Saposnik, 1968, p. 185)

In summary, the exercise of proving local stability under a variety of different restrictions on the matrix JZ(p*) constituted a successful robustness analysis on the dynamic Walrasian model and increased economists' confidence in, and the credibility of, the underlying Walrasian model. But in addition, those assumptions – particularly GS – that had economic interpretations were given special importance and given the choice between two stability conditions, a pure tractability assumption (like QND) or economically interpretable assumption (like GS), more credibility was associated with the latter.

The key papers in the literature on the *global stability analysis* of the dynamic Walrasian model were Arrow and Hurwicz (1958), Arrow, Block, and Hurwicz (1959), and McKenzie (1960). These papers introduced the Lyapunov technique

---

[23] "There seem to be no examples in the literature of utility and production functions which yield diagonal dominance other than of course the GS case" (Hahn, 1982, p. 759)

into the Walrasian literature and it became the standard tool for global stability analysis. Although global stability is a much more powerful result than local stability – proving that the price path converges to p* for any initial price vector is a much more general result than only proving that it converges for initial prices within some epsilon neighborhood – but for our purposes here the story is much the same. Global stability analysis was derivational robustness analysis in precisely the same way as local stability analysis. The mathematical generality and tools were different, but the basic technique of showing that particular additional restrictions imposed on excess demand functions were sufficient for stability was exactly the same (although now the restrictions were assumed to hold on the functions throughout their domain and not just on the Jacobian matrix at p*). And in addition, and perhaps not surprisingly, the assumptions employed were much the same as well: GS, QDD, QND, etc. Global stability was proven for a version of QND (Arrow and Hurwicz, 1958, p. 536), for GS (Arrow, Block, and Hurwicz, 1959, p. 95), and for QDD (ibid., p. 105). Just as with local stability, the literature on global stability was an example of successful derivational robustness analysis, with GS adding a bit more credibility because of its economic interpretation.

The bottom line for this first period of stability analysis in the dynamic Walrasian model is that it ended on a rather rosy note. The robustness analysis associated with the various local and global stability results were generally quite positive. As Arrow and Hurwicz summarized their results: "The nature of our findings can be summarized very simply by saying that in none of the cases studied have we found the system to be unstable under the (perfectly competitive) adjustment process" (1958, p. 529).

3b. Doubts About Robustness of Stability Appear 1960-1970

Counterexamples are to robustness analysis as falsifications are to empirical testing. One way to think about the problem of induction is view it as a stopping problem in empirical testing; each empirical confirmation adds psychological confidence and epistemological justification, but how many confirmations are enough? So too for robustness; each successful exercise in robustness analysis adds credibility and epistemic import, but how many different auxiliary assumptions that give the same result are enough? In the case of empirical testing, a falsification is a definitive stopping rule; not definitive that the hypothesis is false of course, because of Duhem-Quine, but definitive in the sense that it is time to think seriously about the various parts of the relevant test system. Counterexamples to results deduced from abstract mathematical models also serve as definitive stopping rules; not definitive that the model is necessarily flawed, but definitive in the sense that it is time to think seriously about the relationships among the various parts of the model.

Two counterexamples to stability – one by David Gale (1963) and one by Herbert Scarf (1960) – clearly challenged the previous robustness analysis of the dynamic Walrasian model. Neither one of these counterexamples demonstrated any mistakes – errors in mathematics etc. – in the previous results. It was the relevance of these existing results to the credibility of the Walrasian model that started to be questioned.

The Gale counterexample involved a Walrasian economy where at least one good was Giffen. Scarf's main result was a Walrasian economy where goods were perfect complements. Both authors demonstrated that an unstable equilibrium was possible under these special assumptions. Of course these assumptions are inconsistent with all of the results discussed in the previous section (GS, QDD, QND, etc.), but – and this is important for the impact of these counterexamples – these special cases *were consistent* with the underlying assumptions of the basic Walrasian model. Remember, the question was how positive robustness tests on stability of (T) could increase the credibility of basic Walrasian model *as* a model of a perfectly competitive economy. Positive results thus required, i) stability, and ii) consistency with the underlying assumptions of the Walrasian model. The counterexamples of Gale and Scarf were unstable models that satisfied ii). Of course there was no empirical evidence that Giffen goods or perfect complements were common, but they are assumptions that have an economic interpretation and are consistent with the underlying Walrasian model, and therefore stability was not as robust a feature of dynamic Walrasian models as the earlier results suggested. By 1971 Arrow and Hahn would have a very pessimistic tone at the end of their stability chapter:

> There is a distressingly anecdotal air about our investigation; case succeeds case, but it was not found possible to lay down any general principles … At the moment the main justification for the chapter is that there are results to report on the tâtonnement while there are no results to report on what most economists would agree to be more realistic constructions. (Arrow and Hahn, 1971, pp. 321-22)

But the situation was about to get worse.

3c. <u>Robustness Impossibility 1970-1975</u>

The research that turned economists away from the entire project of analyzing the stability of the dynamic Walrasian model was a series of papers in the early

1970s that came to be called the Sonnenschein-Mantel-Debreu (SMD) results.[24] These papers demonstrated that the standard Walrasian assumptions on utility maximizing agents impose almost no restrictions – only continuity and (W) – on market excess demand functions. This has at least two negative implications for the robustness-based stability results on the Walrasian model. On one hand, this means that the assumptions that were shown to be sufficient for stability – GS, QDD, QND, etc. – were *very special* assumptions and would need to be imposed in some sense *from outside* the basic Walrasian model. But how could demonstrating stability under such externally imposed assumptions increase the credibility of the Walrasian model? Secondly, it made counterexamples very easy to produce. If excess demand functions have very little structure, then it is easy to construct examples that are consistent with that structure and yet generate instability. The Gale and Scarf counterexamples were not special cases, they were just the tip of the iceberg. The SMD results were essentially a *derivational robustness impossibility result*. They made it impossible to keep expanding the list of restrictions that implied stability and opened up the floodgates for restrictions that would do just the opposite. Economic theorists had hoped that the core assumptions of the Walrasian model would ultimately prove to be sufficient for stability, but with the failure of robustness, it became quite clear that was not the case.

Early results from robustness analysis had increased economists' confidence in the Walrasian model, but by the late 1970s it was clear that such analysis was a failure and the dynamic Walrasian model needed to be carefully reexamined. The SMD results thus "became a sort of *leitmotiv* (or nightmare) running throughout all research into … stability" (Ingrao and Israel, 1990, p. 317) and this robustness failure exposed a fundamental tension within the core Walrasian model.

> Scarf's paper of 1960 had the merit of pointing out a crucial problem, and Sonnenschein's of 1972 that of showing just why the problem was so hard to solve. The great contradiction revealed is as follows: one of the theory's greatest strengths – its claim to deduce significant results from very general hypothesis about the behavior of economic agents – turns out to be its greatest weakness. The lack of specification of the basic functions … makes it impossible to obtain any significant results at all unless the stopgap measure of imposing very restrictive conditions on the aggregate functions is used. (ibid., p. 346)

[24] The most important papers were Debreu (1974), Mantel (1974), and Sonnenschein (1972, 1973). See Shafer and Sonnenschein (1982) for a survey, and Hands (2012), Ingrao and Israel (1990), and Rizvi (2006) for historical discussions.

A serious discussion of how the economics profession responded to the failure of stability theory is beyond the scope of this paper, but perhaps a few brief comments are in order. As one might expect with a criticism like this there were a variety of different responses. On one hand, there were several responses that seemed to have little impact on the overall credibility of the Walrasian model. These include: the regular economies literature (Debreu 1970); the literature on Non-Walrasian adjustment or disequilibrium trading models (see note 16); new ways of specifying the adjustment mechanism itself that guaranteed stability even with very general excess demand functions (e.g. Smale 1976); and attempts to use information other than direct restrictions on excess demand functions, such as the distribution of income, to guarantee stability (e.g. Hildenbrand, 1994). On the other hand, it seems like the two most enduring responses have been: i) to slowly move away from Walrasian economics: not by explicitly rejecting it, but by reframing it as a special case of non-cooperative game theory, and ii) to put the onus on (T) rather than the Walrasian model, i.e. to drop any kind of a price adjustment mechanism and make general equilibrium simply a *consistency condition* on Walrasian models (so the model is always in equilibrium); this can be done either directly, or by the use of the representative agent. Response (i) has been most common among microeconomists and applied economists, while (ii) has been the primary response among macroeconomic theorists (e.g. Lucas, 1989; Woodford, 2003).

4. Conclusion

This paper supports the literature which argues that derivational robustness can have epistemic import in abstract economic models by examining a particular example from mathematical economic theory: the dynamic Walrasian model. It was argued that derivational robustness first increased economists' confidence in the Walrasian model and then later undermined confidence in the same model. The example demonstrates that derivational robustness *describes the practices of a particular group of very influential mathematical economic theorists* and also supports the philosophical literature that offers *an epistemic justification of such practices* (although a weaker justification than would be provided by traditional empirical confirmation). During the course of the discussion an additional feature of derivational robustness in economic models was introduced – the *economic interpretability* of certain tractability assumptions – and the important role it played in stability analysis was demonstrated. The paper also introduced the notion of the *nesting* of abstract models where the credibility of a specific mathematical model (in this case the Walrasian general equilibrium model) is assessed on the basis of the credibility of its relationship to a higher level, more conceptual, model (the perfectly competitive economy) rather than on the credibility of its relationship to any real economy.

References

Arrow, Kenneth J. (1959), "Toward a Theory of Price Adjustment." In *The Allocation of Economic Resources: Essays in Honor of Bernard Francis Haley*. Moses Abramovitz et al. (eds.), Stanford: Stanford University Press, 41-51.

Arrow, Kenneth J., Block, H. David, and Hurwicz, Leonid (1959), "On the Stability of the Competitive Equilibrium II," *Econometrica*, 27, 82-109.

Arrow, Kenneth and Debreu, Gerard (1954), "Existence of an Equilibrium for a Competitive Economy," *Econometrica*, 22, 265-290.

Arrow, Kenneth J. and Hahn, Frank H. (1971), *General Competitive Analysis*. San Francisco: Holden-Day.

Arrow, Kenneth J. and Hurwicz, Leonid (1958), "On the Stability of the Competitive Equilibrium I," *Econometrica*, 26, 522-552.

Aydinonat, N. Emrah (2007), "Models, Conjectures and Explanations: An Analysis of Schelling's Checkerboard Model of Residential Segregation," *Journal of Economic Methodology*, 14, 429-454.

Aydinonat, E. Emrah and Ylikoski, Petri (2014), "Understanding with Theoretical Models," *Journal of Economic Methodology*, 21, 19-36.

Boumans, Marcel (2005), *How Economists Model the World to Numbers*. London: Routledge.

Cartwright, Nancy (1991), "Replicability, Reproducibility, and Robustness: Comments on Harry Collins," *History of Political Economy*, 23, 143-155.

Cartwright, Nancy (2007), *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.

Cartwright, Nancy (2009), "If No Capacities Then No Credible Worlds, But Can Models Reveal Capacities?," *Erkenntnis*, 70, 45-58.

Debreu, Gerard (1959), *Theory of Value*. New York: Wiley.

Debreu, Gerard (1970), "Economies with a Finite Set of Equilibria," *Econometrica*, 38, 387-392.

Debreu, Gerard (1974), "Excess Demand Functions." *Journal of Mathematical Economics*, 1, 15-21.

de Donato Rodríguez and Zamora Bonilla, Jesús (2009), "Credibility, Idealization, and Model Building: An Inferential Approach," *Erkenntnis*, 70, 101-118.

Elliott-Graves, Alkistis and Weisberg, Michael (2014), "Idealization," *Philosophy Compass*, 9, 176-185.

Gale, David (1963), "A Note on Global Instability of Competitive Equilibrium," *Naval Research Logistics Quarterly*, 10, 81-87.

Gibbard, Allan and Varian, Hal R. (1978), "Economic Models," *Journal of Philosophy*, 75, 664-677.

Grüne-Yanoff, Till (2009a), "Preface to 'Economic Models as Credible Worlds or as Isolating Tools," *Erkenntnis*, 70, 1-2.

Grüne-Yanoff, Till (2009b), "Learning from Minimal Models," *Erkenntnis*, 70, 81-99.

Grüne-Yanoff, Till (2011), "Isolation is Not Characteristic of Models," *International Studies in the Philosophy of Science*, 25, 119-137.

Grüne-Yanoff, Till (2013), "Appraising Models Nonrepresentationally," *Philosophy of Science*, 80, 850-861.

Guala, Francesco and Salanti, Andrea (2003), "Model Robustness in the 'Old' and 'New' Growth Theory," in *Old and New Growth Theories: An Assessment*. N. Salvadori (ed.), Cheltenham: Edward Elgar, 115-126.

Hacking, Ian (1979), "Imre Lakatos's Philosophy of Science," *British Journal for the Philosophy of Science*, 30, 381-410.

Hahn, Frank H. (1958), "Gross Substitutes and the Dynamic Stability of General Equilibrium," *Econometrica*, 26, 169-170.

Hahn, Frank H. (1982), "Stability," in *Handbook of Mathematical Economics*, Vol. II, K. J. Arrow and M. D. Intriligator (eds.), Amsterdam: North-Holland, 745-793.
.
Hamminga, Bert (1983), *Neoclassical Theory Structure and Theory Development*. Berlin: Springer-Verlag.

Hamminga, Bert and De Marchi, Neil (1994), "Idealization and the Defense of Economics: Notes Toward a History," in *Idealization VI: Idealization in Economics*, B. Hamminga and N. De Marchi (eds.), Amsterdam: Rodopi, 11-40.

Hands, D. Wade (2012), "What a Difference a Sum (∑) Makes: Success and Failure in the Rationalization of Demand,: *Journal of the History of Economic Thought*, 34, 379-396.

Hands, D. Wade and Mirowski, Philip (1998), "Harold Hotelling and the Neoclassical Dream," in *Economics and Methodology: Crossing Boundaries*, R. Backhouse, D. Hausman, U. Mäki, and A. Salanti (eds.), London: Macmillan, 322-397.

Hicks, John R. (1939), *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford: Oxford University Press.

Hildenbrand, Werner (1994), *Market Demand: Theory and Empirical Evidence*. Princeton: Princeton University Press.

Ingrao, Bruna and Israel, Giorgio (1990), *The Invisible Hand: Economic Equilibrium in the History of Science*. Cambridge, MA: MIT Press.

Knuuttila, Tarja (2005), "Models, Representation, and Mediation," *Philosophy of Science*, 72, 1260-1271.

Knuuttila, Tarja (2009), "Isolating Representations Versus Credible Constructions? Economic Modeling in Theory and Practice," *Erkenntnis*, 70, 59-80.

Knuuttila, Tarja (2011), "Modelling and Representing: An Artefactual Appproach to Model-Based Representation," *Studies in History and Philosophy of Science*, 42, 262-271.

Kuorikoski, Jaakko; Lehtinen, Aki; and Marchionni, Caterina (2010), "Economic Modelling as Robustness Analysis," *British Journal for the Philosophy of Science*, 61, 541-567.

Kuorikoski, Jaakko; Lehtinen, Aki; and Marchionni, Caterina (2012), "Robustness Analysis Disclaimer: Please Read Manual Before Use!," *Biology and Philosophy*, 61, 891-902.

Kuorikoski, Jaakko and Lehtinen, Aki (2009), "Incredible Worlds, Credible Results," *Erkenntnis*, 119-131.

Lange, Oscar (1944), *Price Flexibility and Employment*. Bloomington, IN: The Principia Press.

Lehtinen, Aki and Kuorikoski, Jaakko (2007), "Unrealistic Assumptions in Rational Choice Theory," *Philosophy of the Social Sciences,* 27, 891-902.

Lehtinen, Aki and Marchionni, Caterina (2011), "Generality in Economic Models," Working Paper.

Lucas, Robert E. (1989), *Studies in Business-Cycle Theory*. Cambridge: MIT Press.

Mäki, Uskali (1994), "Isolation, Idealization and Truth In Economics," in *Idealization VI: Idealization in Economics*. B. Hamminga and N. De Marchi (eds.), 147-168 [*Poznan Studies in the Philosophy of Science and the Humanities*, 38].

Mäki, Uskali (2001), "The Way the World Works (www): Towards an Ontology of Theory Choice," in *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press, 369-389.

Mäki, Uskali (2005), "Models are Experiments, Experiments are Models," *Journal of Economic Methodology*, 12, 303-315.

Mäki, Uskali (2009a). "Realistic Realism About Unrealistic Models," in *Oxford Handbook of the Philosophy of Economics*. H. Kincaid and D. Ross (eds.), Oxford: Oxford University Press, 68-98.

Mäki, Uskali (2009b), "MISSing the World. Models as Isolations and Credible Surrogate Systems," *Erkenntnis*, 70, 29-43.

Mäki, Uskali (2013), "Contested Modeling: The Case of Economics," In *Models, Simulations, and the Reduction of Complexity*, U. Gähde, S. Hartmann, and J. H. Wolf (eds.), Berlin: Walter DeGruyter, 87-116.

Mantel, Rolf R. (1974), "On the Characterization of Aggregate Excess Demand," *Journal of Economic Theory*, 7, 348-353.

McKenzie, Lionel W. (1960), "Stability of Equilibrium and the Value of Positive Excess Demand," *Econometrica*, 28, 606-17

McKenzie, Lionel W. (2002), *Classical General Equilibrium Theory*. Cambridge, MA: MIT Press.

Morgan, Mary S. (2001), "Models, Stories and the Economic World," *Journal of Economic Methodology*, 8, 361-384.

Morgan, Mary S. (2012), *The World in the Model: How Economists Work and Think*. Cambridge: Cambridge University Press.

Morgan, Mary S. and Knuuttila, Tarja (2012), "Models and Modelling in Economics," *Handbook of the Philosophy of Science*, *Vol. 13, Philosophy of Economics*, U. Mäki (ed.), Amsterdam: Elsevier, 49-87.

Negishi, Takashi (1958), "A Note on the Stability of an Economy Where All Goods are Gross Substitutes," *Econometrica*, 26, 445-447.

Newman, Peter (1959), "Some Notes on Stability Conditions," *Review of Economic Studies*, 26, 1-9.

Odenbaugh, Jay and Alexandrova, Anna (2011), "Buyer Beware: Robustness Analysis in Economics and Biology," *Biology and Philosophy*, 26, 757-771.

Patinkin, Don (1965), *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, 2nd edition, New York: Harper & Row [1st edition 1956].

Quirk, James and Saposnik, Rubin (1968), *Introduction to General Equilibrium Theory and Welfare Economics*. New York: McGraw Hill.

Reiss, Julian (2012), "The Explanatory Paradox," *Journal of Economic Methodology*, 19, 43-62.

Rizvi, S. Abu Turab (2006), "The Sonnenschein-Mantel-Debreu Results After Thirty Years," in *Agreement on Demand: Consumer Theory in the Twentieth Century*, P. Mirowski and D. W. Hands (eds.), Durham, NC: Duke University Press [Annual Supplement to Volume 38 of *History of Political Economy*], 228-45.

Samuelson, Paul A. (1941), "The Stability of Equilibrium: Comparative Statics and Dynamics." *Econometrica*, 9, 97-120.

Samuelson, Paul A. (1942), "The Stability of Equilibrium: Linear and Non-Linear Systems," *Econometrica*, 10, 1-25.

Samuelson, Paul A. (1947), *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

Scarf, Herbert (1960), "Some Examples of Global Instability of the Competitive Equilibrium," *International Economic Review*, 1, 157-72.

Schelling, Thomas (1978), *Micromotives and Macrobehavior*. New York: Norton.

Shafer, Wayne and Sonnenschein, Hugo (1982), "Market Excess Demand Functions," in *Handbook of Mathematical Economics*, Vol. II, K. J. Arrow and M. D. Intriligator (eds.), Amsterdam: North-Holland, 671-93.

Smale, Stephen (1976), "A Convergent Process of Price Adjustment as Global Newton Methods," *Journal of Mathematical Economics*, 3, 211-226.

Sonnenschein, Hugo (1972), "Market Excess Demand Functions," *Econometrica*, 40, 549-63.

Sonnenschein, Hugo (1973), "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?," *Journal of Economic Theory*, 6, 345-54.

Sugden, Robert (2000), "Credible Worlds: The Status of Theoretical Models in Economics," *Journal of Economic Methodology*, 7, 1-31.

Sugden, Robert (2009), "Credible Worlds, Capacities and Mechanisms," *Erkenntnis*, 70, 3-27.

Suppes, Patrick (1962), "Models of Data," in *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*. E. Nagel, P. Suppes, and A. Tarski (eds.), Stanford, CA: Stanford University Press, 252-261.

Suppes, Patrick (1967), "What is a Scientific Theory?" in *Philosophy of Science Today*. S. Morgenbesser (ed.), New York: Basic Books, 55-67.

Takayama, Akira (1974), *Mathematical Economics*. Hinsdale, IL: The Dryden Press.

Uzawa, Hirofumi (1960), "Walras' Tâtonnement in the Theory of Exchange," *Review of Economic Studies*, 27, 182-194.

Walras, Léon (1954), *Elements of Pure Economics*, translated by W. Jaffé from the 4th definitive edition, 1926, Homewood, IL: Richard D. Irwin [first edition 1874, Augustus M. Kelley reprint 1977].

Weisberg, Michael (2007), "Three Kinds of Idealization," *Journal of Philosophy*, 104. 639-659.

Weisberg, Michael (2013), *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Woodford, Michael (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

Woodward, James (2006), "Some Varieties of Robustness," *The Journal of Economic Methodology*, 13, 219-240.