

## Causal theories, models and evidence in economics – some reflections from biology

**Michael Joffe**

Imperial College London

One important focus of economic methodology has been on models, e.g. the extent to which they are able to provide explanations (Reiss 2012 and ensuing discussion in *JEM*). This approach fits well with the idea that philosophical analysis should begin with the practice of scientists, or here specifically with economists.

This paper takes a different, complementary, approach. It starts from a cross-disciplinary, comparative perspective on theory development more broadly. Following Marshall's famous dictum, it focuses on biology which, like economics, studies a highly complex reality. The aim is to propose a concept of theory, and of theory development, distinct from models; and also to discuss the relationship between theory in this sense and modelling.

In brief, such a theory seeks to understand a phenomenon or group of phenomena by combining a core set of concepts relating to the causal mechanism with a broad range of evidence. The latter includes evidence on the proposed mechanism, and of the "difference making" type. Such a theory evolves over time, and is likely modified to accommodate new findings as they arise. It is broad and descriptive. Examples include the germ theory of disease, and numerous specific theories about physiological systems. Similar theories occur outside biology too, e.g. tectonic plate theory. They are phrased in the language of causation, not of mathematics; but mathematical models can be developed that relate to specific components of the theory, for example epidemiological models of infectious diseases in populations (different for different diseases), and models of physiological processes such as the biophysics of the circulation. Thus, models are nested within the broader causal theory; this crucially means that scientists are aware of what is omitted in such models.

Such theories do exist in economics, as in recent work on the nature and origin of money (Ryan-Collins et al 2012; Bank of England 2014), and the rise of the modern state (Lindert 2004). However, they are not typical of economic practice. Before the crisis of 2007-08, macroeconomic models notoriously assumed a perfectly operating financial system, without frictions. Problems originating in the financial sector were therefore invisible both to modellers and users. Prevailing practice is to review available relevant models and choose the most suitable, seen as a matter of judgement. Yet without an overarching theory that describes the overall situation causally, the model and its accompanying story define the scope covered by the theory, so it is easy to ignore important omissions.

Relatedly, the term "theory" is typically equated with models in economics. It tends to be based on axioms and assumptions, rather than on evidence as in e.g. biology – which is odd, given the sophisticated and effective methods for causal inference and for testing predictions against evidence used in economics research. These issues are critically discussed in the light of the literature, including the influence of Friedman (1953), notably his two biological examples: the distribution of leaves on a tree, and the billiard player.

**Key words:** economic methodology, economic models, modeling, theory, causal theory, evidence

**JEL codes:** B41 Economic Methodology

## EARLY DRAFT – PLEASE DO NOT QUOTE OR CITE

### 1. Introduction

One important focus of economic methodology has been the analysis of models, e.g. whether they contain – or can in principle contain – true statements about the real world (Mäki 2011), and the extent to which they are able to provide explanations (Reiss 2012 and the ensuing discussion in *JEM* 2013). This approach fits well with the idea that philosophical analysis should begin with the practice of scientists, or here specifically with economists, because modelling is central to the discipline of economics as actually practiced.<sup>i</sup>

This paper takes a different, complementary, approach. It examines the practice of natural scientists who investigate a complicated, heterogeneous and open-ended reality, much like the economy. The main focus is on biology, but the case is made that a similar approach is widely adopted within the natural sciences and is not peculiar to biology. The purpose is to draw methodological parallels between these scientific practices and economics, specifically to argue that they could provide a useful guide for generating reliable knowledge about the economy.

The use of biology as a source of methodological insight accords with Marshall's famous dictum that "The Mecca of the economist lies in economic biology" (Marshall 1920, preface to the 8<sup>th</sup> edition, page xii); this was on the basis that "economics, like biology, deals with a matter, of which the inner nature and constitution, as well as the outer form, are constantly changing" (Marshall 1920, appendix C, page 637). This position is reinforced by the observation that the biological sciences have progressed enormously since Marshall's day.

Opinion is divided as to the degree of success that Marshall himself achieved in following this advice. For example, Gee considered that these statements were "not mere window dressing; they reflect an attitude which profoundly influences the style of analysis throughout the *Principles*" (Gee 1983). On the other hand, according to Hodgson (1993), the biological aspects of Marshall's thought were never well developed by him, and were obliterated by his followers. Nevertheless, a similar opinion to Marshall's was put forward much more recently by Hahn (1991) "I am pretty certain that the following prediction will prove to be correct: theorising of the 'pure' sort will become both less enjoyable and less and less possible ... rather radical changes in questions and methods are required ... the signs are that the subject will return to its Marshallian affinities to biology".

My attempt to draw methodological lessons for economics from biology is distinct from the various attempts to draw parallels between them, e.g. on the grounds that economics is part of biology because it concerns human beings, or more realistically, that there are similar causal processes at work in the two domains. For example, a strand in evolutionary economics seeks concepts in the economic sphere, such as habit or routine on the one hand and innovation on the other, that are considered similar to replication and mutation respectively in the genetic mechanism underlying biological evolution; a further development along similar lines is universal or generalised Darwinism (Hodgson and Knudsen 2013).

My aim is to describe an approach to theory development that is unusual in economics, but not entirely absent from it. The meaning I attach to the term "theory" will emerge from brief

outlines of case studies that convey how natural scientists have gone about their work, rather than attempting a systematic analysis based on the verbal dissection of the various possible meanings of the term as used by methodologists and by scientists. After presenting the case studies, I identify the main features that could be useful in economics, and discuss how theory in this sense relates to models. I then present some examples of the use of a similar approach by economists, and other examples where economic theory could benefit from it. Finally, I make the case that it could be used far more systematically in the development of economic theory than has traditionally been the case, and that this would have the advantage of producing theoretical accounts that bear a systematic resemblance to actual reality.

This cross-disciplinary, comparative perspective is complementary to the existing literature. It is bottom-up in that it focuses on the actual practices of scientists, but with the unusual feature that the scientists in question are biologists and other natural scientists, whereas the target discipline is economics. In principle it is potentially applicable to all branches of economics, unlike universal Darwinism and related approaches which are specific to the more “evolutionary” aspects of economics.

## **2. Broad, empirically-informed causal theories**

The case studies in this section are deliberately chosen so as not to require expert knowledge in biology or other natural sciences. They are intended to be reasonably representative of this methodological approach to doing science, although this cannot be guaranteed – others may judge that I have been unwittingly selective.

### *Germ theory of disease*

The germ theory of disease emerged gradually in western Europe between the 16<sup>th</sup> and 19<sup>th</sup> centuries, at which time it became the established explanation of many diseases. It appears that the essential principles of the theory had long been understood by Ayurvedic medicine in India. Previously, European medicine had been dominated by the miasma theory – the idea that disease was due to a poisonous vapour or mist, emanating from decomposed matter. Related to this, certain types of creature such as maggots were thought to arise from dead flesh by a process known as “spontaneous generation” – rather than from parents, i.e. causal agents very much like themselves.

In 1546, Fracastoro proposed that epidemic diseases are caused by transferable seed-like entities, but at the time this suggestion lacked direct evidence. It also lacked plausibility until micro-organisms were first observed by van Leeuwenhoek, the pioneer of microscopy and microbiology, in the 1640s. In the early 18<sup>th</sup> century, Fracastoro’s idea was linked with the existence of micro-organisms on the grounds that they could provide a plausible causal mechanism, e.g. by Andry and by Bradley. Direct evidence was not provided until Bassi’s demonstration in 1835 that a disease which was decimating the silkworm industry in France and Italy was due to a contagious living entity (now known to be a fungus named *Beauveria bassiana*). Bassi also introduced successful interventions, including the use of disinfectants, separation of the rows of feeding caterpillars, isolating and destroying infected caterpillars, and keeping farms clean.

In the mid-19<sup>th</sup> century, evidence on the transmissibility of certain diseases was obtained – i.e. evidence of the “difference-making” type relating to the phenomenon itself, rather than on

the mechanism. In 1847, Semmelweis realised that puerperal fever was being spread to women in labour by doctors who had been conducting autopsies. Enforcing a hand-washing rule using chlorinated lime water reduced the mortality from childbirth from 18% to 2.2% – although his findings had little impact on the medical profession at the time.

During the 1850s, Snow established that the risk of cholera was related to the household's water supply. In one study, a simple statistical analysis showed that households supplied by the Southwark and Vauxhall Company had more than nine times the risk of those supplied by the Lambeth Company; the latter drew its water upstream from the Thames, before it reached London and its sewage, whereas the former supplied water contaminated by sewage. Another study demonstrated that a severe outbreak in Soho, central London, was spatially clustered around a particular pump – later found to have become contaminated by a cesspit. Snow had no knowledge of the rudimentary germ theory of disease outlined above, but his scepticism of the miasma theory motivated his research – the findings of which were not accepted by government officials until much later.

In subsequent decades, work by Pasteur, Koch and others established beyond any doubt that some diseases are caused by micro-organisms. Their work included research on rabies, anthrax, tuberculosis, and many others. A pattern was emerging: the hypothesis that many diseases are caused by micro-organisms was established, together with increasing detail on related ideas such as immunity. Subsequently, further research showed that some diseases involve also a vector such as the mosquito, e.g. Ross's work on malaria.

How then can we characterise the key features of the germ theory of disease? The core theoretical concepts are that (i) micro-organisms can invade the body and cause disease; and (ii) micro-organisms multiply. A corollary is that such diseases can be spread either directly from person to person, via an inanimate medium such as water, or by a vector. This clearly meets the criterion for “a genuine scientific theory” set out by Reiss (2011): “a small number of explanatory hypotheses that can be used over and over again in the explanation of a wide range of phenomena”.

An important feature of the germ theory of disease is that its language refers to *capacities* of the important entities: the capacity of particular types of micro-organisms to invade the body and cause disease; their capacity for reproduction. Other components also involve a similar language, e.g. the capacity of the human immune system to defend against infection, and the capacity of insect vectors such as mosquitoes to play their role in the micro-organism's life cycle. The theory sets out to describe how (this aspect of) the world works, in terms of the structure and capacities of the component entities, i.e. of causation. It is therefore ontic not epistemic in intention.

As would be expected of a hypothesis involving causation, the history of the germ theory of disease involves the interweaving of different types of evidence – and specifically, both of mechanistic and of difference-making kinds. Moreover, the participating scientists expect that they should eventually fit together neatly, in accordance with the view that scientists operate according to the principle that a causal relationship is one that has a *mechanism* that by its operation *makes a difference* (Joffe 2013).

Another feature of the germ theory of disease is that it evolved over time. The core concepts, from Fracastoro's transferable seed-like entities onwards, emerged long before the evidence was adduced that showed the theory to be true of many diseases. The theory was modified

and amplified as evidence accumulated, e.g. to include insect vectors. A theory of this kind is therefore broad and descriptive, built on an interlocking nexus of evidence, and indefinitely alterable and extendable to new phenomena in the light of new evidence.

It also encompasses heterogeneity: the biology of rabies, anthrax, tuberculosis and malaria are quite different, but the core concepts remain applicable – with the details depending on the specific evidence on each disease.

In addition, such a scientific theory is “allowed” to be incomplete. Not all diseases are caused by infectious agents, and not all that are so caused are contagious. Conversely, not all micro-organisms cause disease – indeed it is now known not only that most are harmless, but that some types of micro-organisms are beneficial, and even essential. Finally, multiple causation is permitted: the impact of an invading horde of micro-organisms may depend on unrelated factors such as nutritional status or genetics.

Several features that apply to the germ theory of disease – and, as I will argue, to scientific theories more generally – do not apply to models; this theme will be discussed in more detail in section 4. These include the causal language of capacities; the ontic rather than epistemic focus; the importance of the role of evidence, of multiple interlocking kinds; and the property that they are broad, descriptive and heterogeneous, as well as being incomplete and allowing multiple causation as described in the previous paragraph.

Models do however appear in the germ theory of disease literature. A prime example is Ross’ model developed in the 1910s, which he called “*a priori* pathometry”: he calculated that below a critical level of the ratio of the number of mosquitoes to the number of humans, a malaria epidemic would be unable to propagate itself. A similar approach, using differential equations to model transmission dynamics, has been extended to cover infectious diseases in general (Anderson & May 1991). These take account of the specific causal properties of each of the diseases.

### *The circulation of the blood in humans*

Before Harvey’s ground-breaking research, published in 1628, it was generally believed that the blood oscillated in the blood vessels and that it was produced in the liver, a theory dating back to Galen in the 2<sup>nd</sup> century and still regarded as authoritative in the 17<sup>th</sup> century. Harvey conducted a large variety of investigations, culminating in his theory that the blood circulates in a closed circuit rather than oscillating, and that it is not newly created in each heartbeat. His research methods included observations and experiments on a large variety of animals, mainly vertebrates but also some invertebrates. For example, by ligating arteries and veins, he established that blood flowed from the heart in the arteries and towards the heart in veins, and that the valves in the veins and in the heart itself guaranteed this. There were two separate circulations, one from the right side of the heart that perfused the lungs, the other from the left side of the heart to the rest of the body. One of his studies was quantitative, and showed that the heart pumps out 540 pounds (245 kilos or about 240 litres) each minute, making the new creation of this quantity of blood highly implausible. As is clear from this description, he combined empirical research with careful attention to mechanistic hypotheses. The heart was merely a pump, not a mystical seat of the spirit.

The main missing element in his account was to explain how the blood got from the arteries to the veins – his idea was that it passed through “pores” in the flesh. This was shown to be

incorrect once microscopes became available, when Malpighi observed capillary vessels in 1660, thus filling the gap.

This particular example of a physiological theory is unusual in that so much was achieved so early by one person. It has not only stood the test of time, but has also been augmented by a large variety of more recent findings – this is far more typical: the growth of this type of knowledge is cumulative. For example, plasma proteins exert a form of osmotic pressure, thus pulling fluid into the circulatory system, balancing the hydrostatic pressure that results from the pumping of the heart. This can be expressed as the Starling equation (1896), which models fluid movement between compartments.

Another important addition to knowledge about the circulatory system is that the force of contraction of the heart muscle increases with a larger volume of blood filling the heart when all other factors remain constant (known by various names such as “Starling’s Law”). This synchronises the output of the heart with the amount of blood reaching it, among other things. The mechanical and molecular basis of this phenomenon is now well understood. In addition, the output of the heart can be affected by other factors, such as exercise.

Although not generally described as a theory (as is typical in physiology), this account of the blood circulation has similar properties to those listed for the germ theory of disease. Its core concepts are the capacities of its components, this time with a heavy emphasis on structure, e.g. valves, as well as the propensity of heart muscle to contract rhythmically, with additional features in the more recent contributions. These include the importance of flow, as well as capacities, and its causal efficacy, e.g. in a quantitative sense – an aspect of causation that perhaps needs more attention both in philosophy and in economics. As with the germ theory of disease, the development of this theory involved the interweaving of different types of evidence, both of mechanistic and of difference-making kinds.

A difference between this example and the germ theory of disease is that there is far less heterogeneity: one can refer to “the” human blood circulatory system, and the theory was developed on this basis, whereas the germ theory is different for every disease. The reason for this is that it is an evolved system in a single species. Evolution leads to the development of systems that are rather similar across individuals. This is typical in physiology, whereas the germ theory involves the co-evolution of different species, which is far less predictable.

Causal incompleteness and multiple causation also feature, e.g. that “Starling’s Law” only applies to healthy circulatory systems (a *ceteris paribus* condition), and that the output of the heart can be affected by different factors. In addition, important aspects of the physiology of the circulation have been captured as mathematical models such as Starling’s equation; this is *nested within* the broader descriptive, causal theory.

### *Continental drift and tectonic plate theory*

As early as 1596, Ortelius noted the similarity of the coastlines of continents on the opposite sides of the Atlantic Ocean, based on maps of the time. Many others followed, but the theory of continental drift, as it became known, was contradicted in the mid-nineteenth century by the Dana’s Permanence Theory, which maintained that the main outlines of continents had existed since the earliest time.

Further types of evidence accumulated that supported continental drift, notably similar rock formations on the either side of oceans, as well as observations on ancient forms of animal and plant life that similarly were present in widely-separated continents. The evidence for this theory was proposed in a more complete way by Wegener in 1912, but it was almost half a century before it was accepted.

One of the problems was the lack of a plausible mechanism that would be strong enough to move continents. Also, some of the proposed ideas for this required e.g. that the Earth should have been increasing in size.

The situation began to change in the late 1950s, with the arrival of new types of evidence, notably the observation of seafloor spreading. This occurs at mid-ocean ridges, such as the one that runs down the middle of the Atlantic Ocean. Volcanic activity creates new seafloor, as magma rises through fractures and then cools. This suggested that the seafloor itself moves and carries the continents, rather than the continents ploughing through the sea as in earlier theories. This accorded with other evidence that was accumulating, e.g. on patterns in the Earth's magnetic field. It led to the development of tectonic plate theory, which quickly became widely accepted.

The new theory involved observations on the different densities of continental and oceanic parts of the Earth's crust, convection flows in the upper mantle, and the recognition of the different types of boundary that occur where the plates meet. The boundaries of the plates were mapped, with seven or eight major plates and numerous minor ones being identified. It was found that the formation of new crust in the mid-ocean ridges is balanced by the loss of crust into the mantle as part of the process of subduction, where one plate is pushed under another (implying that the Earth does not have to change size).

The principles of plate tectonics were able to explain more than the phenomena that had led to the idea of continental drift: volcanic activity and earthquakes are also attributable to the same forces, and occur particularly at subduction zones. Mountain building is also explicable from the interaction of tectonic plates, when one is crumpled and pushed upwards. However, the exact mechanism involved in plate tectonics is not yet established; the main candidates are mantle convection dynamics, gravitation, and Earth rotation.

The core concepts of the current concept of plate tectonics thus involve magma flows in mid-ocean ridges, differential densities of parts of the Earth's crust, convection flows in the upper mantle, the physical processes of subduction and mountain building, etc. These concern the capacities, including the structure and composition, of the relevant entities, as well as flows. The language is causal.

Once again we see different types of evidence interweaving, with a variety of difference making evidence for the older continental drift idea, some in geology and some in biology – both of living organisms and of fossils. In this story, it was only the recognition of the plausibility of a mechanism that led to its acceptance, and this was based on observations of the ocean floor including ridge spreading and magnetism (“magnetic striping”).

A conspicuous feature of plate tectonic theory is that it is able to explain not only the group of phenomena relevant to continental drift, but also a range of other important topics such as volcanic activity, earthquakes and mountain building. It may not accord perfectly with Reiss'

“small number of explanatory hypotheses”, but the range of explanatory power of a single causal account is nevertheless impressive.

As noted for previous examples, the theory evolved over time, but in this case the history is of a precursor version, continental drift. It set out to explain the same range of phenomena, but was unsuccessful.

In this context, as with the Starling equation that models the balance of forces affecting the fluid in the circulatory system, mathematical models can be developed that relate to specific components of the overall theory. For example, Le Pichon and Sibuet (1981) model the formation of a particular type of ocean and continent boundary (“passive margin”), using a mathematical stretching model, and test this against data from the north-eastern Atlantic. As with the Starling equation, the model is embedded within the broader causal theory, i.e. it is concerned with phenomena within an empirically-based causal account of how this aspect of the world actually works.

### **3. Theories in economics**

Cumulative theory-building as described in the previous section, which is widely used within many of the natural sciences, is not the mainstream methodology within economics. Rather, the usual approach is to go straight for a model. There are exceptions, however.

Recent development of the study of the nature and origin of money is one example of good practice. For some decades, post-Keynesians and others have been emphasising that the assumption of monetary neutrality, which is routine in orthodox economics, obscures many central features of modern economies. The latter also involves the assumption that money is created by the state. There are numerous important issues here that are beyond the scope of the present paper, but the central insight is that money is endogenous, i.e. it is created within the economy in response to particular influences. There has been much progress recently in describing precisely how this works (Ryan-Collins et al 2012). In brief, the vast majority of money is created by banks – 97 percent in modern-day Britain, for example – and this is done when loans are made; the corollary is that repayment of loans involves the destruction of money. The implication is that the money supply depends on the activity of the banking sector, and is a side effect of banks’ attempts to generate profits by making loans – which depends on the demand for loans, as well as on banks’ assessment of their risk. Essentially the same account has now been published by the Bank of England in two important papers (McLeay et al 2014a and 2014b).

This is in marked contrast to the orthodox account, e.g. as it appears in textbooks. Not only is money creation assumed to be a state activity and exogenous (uninfluenced by what is going on in the economy), it also involves a systematic misrepresentation of the banking sector. It is said that banks accept deposits, and then use these “loanable funds” to make loans; there is also a device that explains how banks can lend a multiple of the money that they receive as deposits – the “money multiplier”. This is now accepted to be wholly false, but the notion of loanable funds is still frequently heard in ordinary discourse, and the money multiplier is still being taught.

The way that the correct description of the money generating mechanism was achieved was by the patient documenting of what is actually occurring in the economy, describing what

happens using such basic tools as balance sheets. This is reminiscent of the emphasis on description of the reality in front of the observer that characterises natural sciences such as biology, followed by generalisation and explanation. It should actually be less problematic to do this in economics, for example compared with Harvey's task when trying to describe what was happening in the fast-moving circulatory system, but this would involve overcoming what appears to be researchers' aversion to describing the obvious, i.e. what is widely seen to be happening.

Another example of theory development based on systematic empirical work is a two-volume study of the growth of the modern state (Lindert 2004). This describes the growth of the state qualitatively and quantitatively in each of the major countries that developed rapidly after the industrial revolution, together with an analysis of the causal factors in that country. It then provides an overview of the forces behind state growth, while acknowledging the between-country heterogeneity. Thus it encompasses description, generalisation and explanation, as well as the limits to generalisation imposed by factors specific to each country. This use of comparative economic history is a good model for developing theory, not least because it ensures that any explanation or suggested causal mechanism corresponds to the spatial and temporal patterns that actually occurred, as well as paying attention to specific factors that may have been present in certain countries.

A prominent example of how orthodox theory is seriously problematic relates to the 2008 financial crisis, and especially to the failure of most of the economics profession either to predict it in advance, or to understand it once it had happened. The standard workhorse model in macroeconomics for some time has been the DSGE ("Dynamic Stochastic General Equilibrium") model. In fact there is a large variety of DSGE models, with various features. But in the run-up to the crisis, the models that were in use assumed the absence of financial frictions – this may not always even have been done explicitly; rather, the financial sector and any possible problems arising in it were just absent from the model. Since the crisis, it has proved unproblematic to add the necessary financial component, but this is little consolation. In a wide variety of modelling traditions, one of the basic tenets is that one should be explicit in establishing the boundaries of the model, and clear about what has been left out (see e.g. Sterman 2000). The issue here is, how do the modeller, and the user of the model and its results, realise what has been omitted?

If models are conjured out of the air, as appears often to be the case with economic theory, there is no obvious answer to this question, apart from exhortation to be careful. However, if the approach to theory development advocated in this paper is followed, and models are then nested within a broader, empirically-based causal theory, then it is relatively easy to see what has been omitted in the modelling process.

Examples of orthodox methods of theorising that run into trouble are easy to find. A prime example is the classic paper by Lucas (1990), which attempted to explain why capital does not flow from rich to poor countries – standard theory would predict that all capital would flow from capital-rich to capital-poor countries, as the latter would have higher returns in accordance with diminishing returns. This paper is a gem of its kind, the language and the mathematics both being highly elegant.

Lucas started with neoclassical theory, and then compared reality with it, resulting in a puzzle or "paradox" that he attempted to explain. But let us start at the other end. From the early 1990s until recently, the main observation about international capital flows is not the lack of

massive rich-to-poor country flows, but the presence of massive poor-to-rich country flows. The prime example is the US deficit with China, but the phenomenon is far more widespread.

This is very straightforward to explain, if one combines a variety of types of evidence. The 1978 reforms in China (institutional description), consequent high productivity at very low cost (comparative price data), high profitability and capital accumulation in the corporate, household and public sectors (data on capital) generated a huge surplus above what was required for investment, even with the high investment levels that characterised the Chinese economy. The excess capital, much in hard currency, flooded out of China and was used to purchase US Treasury Bonds, real estate in America and elsewhere, and much else. Thus, an explanation involving large-scale capital generation and consequent flows emerges from the evidence.

China is unusual in the scale of this phenomenon, but in fact the earlier successful East Asian economies had a parallel experience on a smaller scale. Also, a similar phenomenon has been observed with other economies that have generated capital in excess of needs, notably the major oil exporters.

What then is the problem with Lucas' analysis? Part of the answer is that it is myopic: instead of seeing the world as it is, its scope is determined by the dominant theoretical perspective – its starting point is epistemic not ontic. And even though it then confronts the theory with the data, its vision remains restricted. This is reminiscent of the argument that economic analysis should be data-first not theory-first (Juselius 2011).<sup>ii</sup>

One other potential problem can occur when models are not embedded in a wider causal theory. Economists typically describe their situation as requiring familiarity and expertise with an array of different models that cover a particular topic. The task is then to select the most appropriate for a particular purpose, using their accumulated judgement. The issue is the relationship of the content of the available models with substantive knowledge *that is not incorporated in this "theory"*. For example, economists working on Lucas' puzzle must surely know how China (and other similar countries) came to be capital exporters – everyone does. Yet the recent literature ignores the obvious, and seeks to explain how the data can be combined with standard theory. For example, one prominent strand is that the capital-export phenomenon is due to financial weakness (Prasad et al 2007; Cabellero et al 2008; Sandri 2010; Buera and Shin 2011; Gourinchas and Jeanne 2013)! In such a situation, pointing out that capital export from China has an obvious explanation may not be accepted as evidence against the theory, because it is already known to the theorist – there is a danger that the co-existence of substantive knowledge with "theory" that does not incorporate it can actually protect bad theory.

An example where the approach advocated in this paper promises to be fruitful, but where it has been ignored by mainstream economic theory, is the question of differential growth rates. Since the industrial revolution, the "great divergence" led to a gigantic chasm in prosperity opening up between the countries that industrialised early and the rest of the world, subsequently greatly modified by certain other countries experiencing large-scale catch-up growth, but still with very large differentials between the developed world and, for example, most of sub-Saharan Africa (Pomeranz 2000). But why? As Lucas (1988) has said, "The consequences for human welfare involved in questions like these are simply staggering: Once one starts to think about them, it is hard to think about anything else".

The literature on economic growth is vast and heterogeneous, and cannot be summarised here, but the dominant feature is that orthodox theory treats growth as resulting from exogenous technical change.<sup>iii</sup> The standard production function approach is augmented by exogenous productivity growth (Lucas 1988), and complementary to this are historical accounts that describe the inventions, e.g. in the British nineteenth century (Mokyr 2002; Lipsey and Carlaw 2004). This analysis raises the question, if technical change is exogenous, why do particular societies at particular times have an abundance of the relevant type of invention?

Such an account is unable to explain why capitalist economies since the industrial revolution have a unique growth record. As Baumol (2002) noted, “the spectacular and unprecedented growth rates of the industrialized market economies ... set them apart from all alternative economic systems ... growth rates for about one and a half *millennia* before the Industrial Revolution are estimated to have been approximately *zero* ... In contrast, in the past 150 years, per capita incomes in a typical free-market economy have risen by amounts ranging from several hundred to several thousand percent!” [emphasis in the original]. And yet his response is to provide a model based on oligopolistic competition and the importance of research and development, features that empirically do not correspond to the spatial and temporal distribution of this growth.

McCloskey (2010) has gone further, documenting that the magnitude of the transformation of early-industrialising countries is many-fold too large to be explained by the sum total of all conventional explanations. But her positive suggestion, that the emerging industrial middle class was accorded more respect than before, is unconvincing as an explanation of the immense transformation that occurred – and particularly that it occurred in culturally hugely diverse societies such as Victorian Britain, America and Germany in the late nineteenth century, Meiji Japan, the four “little tigers”, and post-reform communist China, yet not in the various other places where respect for business may have been high but the growth rates have not matched this. These two examples demonstrate that the problem may sometimes not be under-determination; rather it is the lack of *any* applicable theory.

A more realistic approach would start from an examination of the places and times where such a transformation occurred (including those just listed), and those where it did not occur, and draw conclusions from the comparative descriptions, possible generalisations and also exceptions, thereby seeking explanations that fit the observations. As far as I am aware, this has rarely been attempted,<sup>iv</sup> and is certainly not part of the discourse of either orthodox or heterodox economics.

#### **4. Theories, models and causation**

Most research in economics nowadays involves empirical work. Good datasets have become available that cover a large variety of topics, and they are constantly improving and being extended. The same applies to economic history, in which the comparative method is also now strongly established. Methods of analysis have improved too, with the development of sophisticated and effective methods for causal inference and for testing predictions against evidence.

It is therefore odd to find a great deal of economic reasoning still starting from “basic theory” that has no empirical foundation, and is really a modelling perspective, methodology and set

of models, rather than theory in the sense used in this paper (and in much of natural science). In fact, in many sub-disciplines the discourse largely concerns evidence. Yet it remains true that the traditional models retain a central place, and accumulating evidence does not always lead to the abandonment of a conventional starting point, even when the two are in conflict.

Rather, we have the accumulation of “puzzles”. The Lucas puzzle, discussed above, is admittedly old now, but for example in labour economics we have the wage flexibility puzzle (Pissarides 2009), the unemployment volatility puzzle (Chodorow-Reich and Karabarbounis 2013), and the unexplained employer size-wage effect (ESWE) (Adamczyk 2015). Modelling in labour economics is still largely based on the assumption of competitive markets, even though this is clearly less appropriate in this context than with many other markets; and matching functions are also widely used. The occasional dissenting voice is heard, e.g. that a better assumption is monopsony or oligopsony (Manning 2003), but even here the starting point is that the traditional competitive-market assumption predicts some odd things that are not observed in reality – it is an incremental change, albeit a potentially important one. In labour economics, as in many other sub-disciplines, there is abundant good evidence, but the link from evidence into theory does not operate – the “theory” remains largely untouched.

An alternative, argued for in this paper, is that different strands of evidence can be combined with each other, and with explanatory ideas, to generate a broad, empirically informed theory. One principal aim is to identify the causal process(es) involved in generating the phenomena. Development of such a theory typically takes place over time, it is not a one-off invention. The same theory may have core concepts that explain a heterogeneous reality – as with the germ theory applied to different diseases, *mutatis mutandis* – which is likely to be important in economics. Other features can include that a theory can be complicated if necessary, and that it can be incomplete in the sense of not explaining everything. The operation of a causal factor may also be strongly modified by other causal influences – multiple causation. Models can then be developed, e.g. mathematical models, nested within such a theory, to simulate, quantify or otherwise illuminate particular aspects.

This would be a natural way of developing conceptual categories that correspond to natural categories (“carving nature at its joints”). It also has a strong ontic emphasis, and a focus on causation. These features are widely agreed to provide a good basis for explanation, even among those who argue that other bases are possible, e.g. that explanations can be epistemic (Reiss 2013). A theory in this sense can also be said to be true or false – or perhaps better, that it is able to possess some degree of truth.

Models do not necessarily possess these characteristics. It is controversial whether or not they can be true or false: Hausman (2013) believes that they cannot be, because they are artificial constructions not sentences, whereas Mäki (2013) considers that they can sometimes be true and/or contain true elements. They are typically not phrased in the language of capacities, even when they accurately capture a causal feature of the target system – rather, they tend to be set out as relationships between variables (at least this is true of economics). Models are constructions, and therefore epistemic (Reiss 2013), whereas empirically informed theories are intended to represent the real world accurately and are thus ontic. And whereas evidence is central to such theories, it does not play an explicit part in model construction – although obviously the modeller typically takes account of the real world when deciding on what the model should contain. The key strength of a model is its simplicity, whereas a theory can be quite an elaborate structure that brings together multiple types of evidence with explanations;

they can also apply to heterogeneous situations, as we saw with the germ theory, and allow for multiple causation.

Among economic methodologists, there appears to be wide agreement that one important feature of a model – possibly the main criterion in judging it – is its resemblance to its target system. Hausman (1992) uses the term “theory” to mean a model plus a theoretical hypothesis, where a theoretical hypothesis is of the form ‘target system *T* is of the kind model *M* defines’. Mäki (2013) also emphasises the importance of models resembling their targets (what he calls the “semantic” issue), adding that this needs to be settled before dealing with epistemic concerns.

Others express this specifically in causal terms. Thus Alexandrova & Northcott (2013) argue that in principle, models can be successful at isolating the capacities that occur in the target system. However, they state that in practice this does not happen in economics (and in some biology). They advocate “good old-fashioned experiment-tested causal hypotheses” and “close empirical study”. Similarly, Rol (2013) praises models that capture something important such as a causal mechanism – although his example is a Madrid metro map, where the resemblance is structural rather than causal. Grüne-Yanoff (2013) agrees about the importance of correspondence, adding that the criterion is the stability of the causal factor across different environments, which is an empirical question.

Thus there appears to be wide agreement that models vary in the degree to which they represent their target system, and that causal properties or capacities are an important aspect of this, possibly the most important. The success of economics in achieving this is judged differently by different authors – but such a judgement inevitably needs to be case by case, so that an overall generalisation is difficult. What does seem to be generally true is that there is no systematic method for ensuring this correspondence within economics.

The classic methodological statement of Friedman (1953) is interesting in this respect. Here I will confine my remarks to his two biological examples, the distribution of leaves on a tree, and the billiard player.

In the latter case, Friedman’s stated aim is to predict the shots, using Newtonian mechanics, even though the player does not actually use Newton’s equations while playing. Friedman considered that this was a good analogy for what economists do when their equations describe economic behaviour in terms that the economic agent would not recognise. The model describes the behaviour “as if” it were a calculation, while acknowledging that the actual neurophysiological processes do not correspond to the description. Thus, this account does not aim at uncovering the causal process involved, as a biologist would do. For example, a neuro-physiologist would seek to explain how expertise in billiard playing is developed and executed in terms of neuronal pathways and neurotransmitters. Scientifically Friedman’s account is impoverished.

In the case of leaves on a tree, it is true that it is not trivial to observe that the leaves are distributed “as if” designed to catch the maximum sunlight.<sup>v</sup> But a biologist’s task is to try and explain this distribution. There is a ready-made explanatory framework in the theory of natural selection: trees with a genetic tendency towards the most beneficial leaf arrangement are more likely than others to survive and pass on their genes. The biologist’s task would be to identify the particular pathways in the general framework, for example, what competitive pressures are acting on the tree, perhaps in terms of climate or of a changing environment;

identifying the development process and relevant genes, establishing the biochemical pathways by which they act; and so on. The idea that an “as if” explanation is satisfactory would be difficult to comprehend.

On Friedman’s view therefore, causal understanding is irrelevant to the economic modeller. To the extent that this has been influential, it may have detracted from the degree to which models in economics represent the capacities in the target system.

One way that a model is connected to (features of) the world is by the story that accompanies the formal part of the model (Morgan 1999), placing a considerable burden on the adequacy of the story. It may be that the perceived problem that much economic theory is unrealistic results from a resulting poor connection, rather than with internal features of the model itself, such as simplifying assumptions. This “semantic” issue needs to be established first, because if it is seriously deficient then it could be argued that other aspects of the model lose their interest, e.g. the other function that Morgan specifies for a model, to demonstrate outcomes within the model.

What economic theorizing needs is a more robust process for ensuring the correspondence of key aspects of each model with the target system, especially its central causal features. One way of doing this would be to emulate the methods of natural sciences such as biology, as outlined above.

What difference would this make? In the case of Lucas’ puzzle of 1990, basing theory on an empirical account would quickly solve the problem – the export of capital from China during its years of mega-profitability is no mystery, and the same applies to other such countries.

To take the more complicated example of capitalist growth, if the starting point were to be comparative economic history, that clearly describes where and when such growth has occurred and where it has not, this would guarantee that putative explanations would at least fit with this macroscopic evidence. Then we might have several rival models and the problem of under-determination, rather than the current situation where it is hard to find any plausible ideas that correspond to the spatial and temporal observations. A model would also need to be able to account for the magnitude of the transformation as well as its functional form which has typically been exponential (suggesting a reinforcing feedback process). Another angle is the nature of micro-foundations: the fact that growth of this type is so limited spatially and temporally indicates that any micro-foundations cannot be based on a notion of universal human nature (e.g. optimization), but rather must rest on a different foundation, e.g. institutions.

## **5. Conclusion**

In the economic methodology literature, models in economics are often compared with models in the natural sciences – typically in physics and mechanics, such as “the idealizations of frictionless plane, perfectly elastic gas molecule, rigid body, planets as mass points, two-body solar system” (Mäki ). This paper has taken a different perspective, which may be useful for a discipline that deals with such a complicated and open-ended reality. The core idea is that “theory” should be regarded as something quite different from a model, a modelling perspective, or a group of models.

Reiss (2011) also distinguishes between models and theory in the context of what economics could learn from other disciplines, giving the examples of Newton, Darwin and Einstein (as well as, less convincingly, Marx and Freud). As previously stated, his concept of a theory is “a small number of explanatory hypotheses that can be used over and over again in the explanation of a wide range of phenomena”. I agree that this is the aim, and is achieved in the theories associated with these great scientists.<sup>vi</sup> In this paper, I have tried to describe one of the major ways that the natural sciences, such as biology, achieve reliable knowledge of a reality in which (again in Marshall’s words) “the inner nature and constitution, as well as the outer form, are constantly changing”. The examples discussed above – e.g. the germ theory of disease and tectonic plate theory, are themselves successful, even if not quite in the league of Newton, Darwin and Einstein. But think about the germ theory in 1850, or the continental drift idea in 1950 before tectonic plate theory – these are more equivalent to the situation in much of economics. The issue is how to get from where we are now to the more desirable situation of having a satisfactory theory.

I suggest that one possible way is to use this approach as an exemplar. It is not the only one, and economists will continue to use a variety of methodological approaches. However, it is probably the one most likely to achieve alignment between theorising and reality.

## References

- Adamczyk, S. (2015), 'The employer-size wage effect: a comparison between Germany and the USA', *The Bonn Journal of Economics*, IV(1), 6-22.
- Aggarwal, R. (2014), 'Animal spirits in financial economics: a review of deviations from economic rationality', *International Review of Financial Analysis*, 32(1): 179-87.  
[doi:10.1016/j.irfa.2013.07.018](https://doi.org/10.1016/j.irfa.2013.07.018).
- Aghion, P. and Howitt, P. (1998), *Endogenous growth theory*, (Cambridge, MA: The MIT Press).
- Alexandrova, A. and Northcott, R., (2013), 'It's just a feeling: why economic models do not explain', *Journal of Economic Methodology*, 20(3), 262-67,  
<http://dx.doi.org/10.1080/1350178X.2013.828873>.
- Anderson, R.M. and May, R.M. (1991), *Infectious diseases of humans: dynamics and control*. (Oxford: Oxford University Press)
- Baumol, W. J. (2002), *The Free-market Innovation Machine*, (Princeton, NJ: Princeton University Press).
- Buera, F.J., Shin, Y. (2011), 'Productivity growth and capital flows: The Dynamics of Reforms', National Bureau of Economic Research Working Paper 15286.
- Caballero, R.J., Farhi, E., Gourinchas, P.-O. (2008), 'An Equilibrium Model of 'Global Imbalances' and Low Interest Rates', *American Economic Review* 98(1), 358-93.
- Chodorow-Reich, G., Karabarbounis, L. (2013), 'The cyclical nature of the opportunity cost of employment', NBER Working Paper 19678, November 2013. Available at <http://www.nber.org/papers/w19678> [accessed 25 December 2015].
- Friedman, M. (1953), 'The methodology of positive economics', reprinted in Hausman, D.M. (ed.). 1994. *The philosophy of economics: an anthology*, 2<sup>nd</sup> edition, (Cambridge, Cambridge University Press).
- Gee, J.M.E. (1983), 'Marshall's views on 'short period' value formation', *History of Political Economy*, 15(2), 181-205.
- Gourinchas, P.-O., Jeanne, O. (2013), 'Capital flows to developing countries: the allocation puzzle', *Review of Economic Studies* 80(4), 1484-1515.
- Grüne-Yanoff, T. (2013), 'Genuineness resolved: a reply to Reiss' purported paradox', *Journal of Economic Methodology*, (20)3, 255-61.
- Hahn, F. (1991), 'The next hundred years', *The Economic Journal*, 101(1), 47-50.
- Hausman, D.M. (1992), *The inexact and separate science of economics*, (Cambridge: Cambridge University Press), section 5.2.

Hausman, D.M. (2013), 'Paradox postponed', *Journal of Economic Methodology*, (20)3, 250-54.

Hodgson, G.M. (1993), 'The Mecca of Alfred Marshall', *Economic Journal*, 103, 406-15.

Hodgson, G.M. and Knudsen, T. (2013), *Darwin's conjecture: the search for general principles of social and economic evolution*. (Chicago: The University of Chicago University Press).

JEM. (2013), *Journal of Economic Methodology*, 20(3), *passim*.

Joffe, M. (2011), 'The root cause of economic growth under capitalism', *Cambridge Journal of Economics* 35, 873-96. doi:10.1093/cje/beq054

Joffe, M. (2013) 'The concept of causation in biology', *Erkenntnis*, 78, 179-97.

Juselius, K. (2011), 'Time to reject the privileging of economic theory over empirical evidence? A reply to Lawson', *Cambridge Journal of Economics*, 2011, 35, 423-36. doi:10.1093/cje/beq024

Le Pichon, X and Sibuet J.-C. (1981), 'Passive margins: a model of formation', *Journal of Geophysical Research*, 86(B5), 3708-20.

Lindert, P. H. (2004), *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century*, (Cambridge, UK: Cambridge University Press).

Lipsey, R. G. and Carlaw, K. I. (2004), 'Total factor productivity and the measurement of technological change', *Canadian Journal of Economics*, 37, 1118-50.

Lucas, R.E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, 22, 3-42.

Lucas, R.E. Jr. (1990), 'Why doesn't capital flow from rich to poor countries?' *American Economic Review*, 80 (2): 92-96.

Mäki, U. (2009), 'Realistic realism about unrealistic models', in Kincaid, H. and Ross, D. (eds.) *Oxford book of the philosophy of economics*, (Oxford University Press).

Mäki, U. (2011), 'Models and the locus of their truth', *Synthese*, 180(1): 47-63.

Mäki, U. (2013), 'On a paradox of truth, or how not to obscure the issue of whether explanatory models can be true', *Journal of Economic Methodology*, (20)3, 268-79.

Mäki, U. (2013), 'On a paradox of truth, or how not to obscure the issue of whether explanatory models can be true', *Journal of Economic Methodology*, (20)3, 268-79.

Manning, A. (2003), *Monopsony in Motion: Imperfect Competition in Labor Markets* (Princeton, NJ: Princeton University Press).

- Marshall, A. (1920), *Principles of Economics*, 8<sup>th</sup> edition, (New York, NY: Cosimo).
- McCloskey, D.N. (2010), *Bourgeois dignity*, (Chicago, IL: University of Chicago Press).
- McLeay, M, Radia, A, Thomas, R. (2014a), 'Money in the modern economy: an introduction', *Bank of England Quarterly Bulletin*, Q1, 1-10.
- McLeay, M, Radia, A, Thomas, R. (2014b), 'Money creation in the modern economy', *Bank of England Quarterly Bulletin*, Q1, 14-27.
- Mokyr, J. (2002), *The Gifts of Athena: Historical Origins of the Knowledge Economy*, (Princeton: NJ, Princeton University Press).
- Moneta, A. and Russo, F. (2014), 'Causal models and evidential pluralism in econometrics', *Journal of Economic Methodology*, 21(1), 54-76, DOI:[10.1080/1350178X.2014.886473](https://doi.org/10.1080/1350178X.2014.886473)
- Morgan, M. (1999), 'Learning from Models', in Morgan M., and Morrison, M., (eds.) *Models as mediators. perspectives on natural and social science*, pages 347–88, (Cambridge: Cambridge University Press).
- Nelson, R.R. and Winter, S.G. (1982), *An evolutionary theory of economic change*, (Cambridge, MA: The Belknap Press of Harvard University Press).
- Pissarides, C.A. (2009), 'The unemployment volatility puzzle: is wage stickiness the answer?', *Econometrica*, 77, 1339-69.
- Pomeranz, K. (2000), *The Great Divergence*, (Princeton, NJ: Princeton University Press).
- Prasad, E, Rajan, R, Subramanian, A. (2007), 'The paradox of capital', *Finance & Development (IMF)* 44(1). Available at <http://www.imf.org/external/pubs/ft/fandd/2007/03/prasad.htm> [accessed 28 December 2015].
- Reiss, J. (2011), 'Theory, generalisations from cases and methodological maxims in evidence-based economics: Responses to the reviews by DiNardo, Guala and Kincaid', *Journal of Economic Methodology*, 18(1), 93-96.
- Reiss, J. (2012), 'The explanation paradox', *Journal of Economic Methodology*, 19(1), 43-62.
- Reiss, J. (2013), 'The explanation paradox redux', *Journal of Economic Methodology*, 20 (3), 280-92. <http://dx.doi.org/10.1080/1350178X.2013.828874>
- Rol, M. (2013), 'Reply to Julian Reiss', *Journal of Economic Methodology*, (20)3, 244-49.
- Ruse, M. (1973), *The philosophy of biology*. (London, Hutchinson & Co Ltd), p195.
- Ryan-Collins, J, Greenham, T, Werner, R, Jackson, A. (2012), *Where does money come from? – a guide to the UK monetary and banking system*. (London, New Economics Foundation).

Sandri, D. (2010), 'Growth and Capital Flows with Risky Entrepreneurship', International Monetary Fund Working Paper WP/10/37.

Spanos, A. (2005). Structural equation modeling, causal inference and statistical adequacy. In P. Hajek, L. Valdes-Villanueva, & D.W. (Eds.), *Logic, methodology and philosophy of science* (pp. 639–661). London: King's College.

Spanos, A. (2006). Where do statistical models come from? Revisiting the problem of specification. *Lecture Notes-Monograph Series*, 49, 98–119.

Spanos, A. (2010). Graphical causal modeling and error statistics. Exchanges with Clark Glymour. In D. G. Mayo & A. Spanos (Eds.), *Error and inference: Recent exchanges on experimental reasoning, reliability and the objectivity and rationality of science* (pp. 364–375). Cambridge: Cambridge University Press.

Sterman, J.D. (2000), *Business dynamics: systems thinking and modelling for a complex world*, (Boston, MA: Irwin McGraw-Hill).

---

## Notes

<sup>i</sup> Throughout this paper, I am taking the term “model” to indicate an *a priori* model, rather than a statistical model, in line with the usage in this literature (usually this point is not made clear). There is also a parallel literature discussing the relationship between statistical and causal models, in which some related themes occur, e.g. that statistical inferences need to be model-specific, implying that statistical models are also different from “theory” in the sense used in this paper (Spanos 2005, 2006, 2010; Moneta & Russo 2014).

<sup>ii</sup> A further point concerns flows: it is well-recognised that neoclassical theory tends to be static. But it is less well understood that this involves a systematic neglect of flows, which can have causal efficacy as in this case. Neglect of flows is deeply rooted in neoclassical methodology: for example in basic consumer theory, rather than the whole situation of the consumer being explored, the analysis fixes the buying power available (the “budget constraint”), so that the theory now becomes reduced to decision theory. The flows become invisible – which is odd, given that market exchange is always and everywhere a two-way flow.

<sup>iii</sup> In addition there are also at least two major “post-Schumpeterian” approaches: endogenous growth theory (Aghion and Howitt 1998) and evolutionary economics (Nelson and Winter 1982). Neither of them focuses on the temporal and spatial distribution of capitalist growth.

<sup>iv</sup> An exception is a paper that seeks to account for these phenomena in terms of an arms race between firms that have specifically capitalist features, which is able to explain why such growth did not occur anywhere in the world before the industrial revolution (Joffe 2011).

<sup>v</sup> Methodologically, Friedman’s statement is functional: the distribution of leaves fulfils the need of the tree to capture maximal sunlight. The problems of such potentially teleological arguments are well recognised by scientists, and in the philosophy of science. Philosophers of biology have long realised that they can readily be overcome *in the biological context*, because it is straightforward to make a “translation of talk of functions into terms of talk of adaptations”, i.e. a causal one based on differential survival and reproduction (Ruse 1973); it then also becomes a historical account in real time. Outside biology such a translation cannot in general be made; a functional statement may not correspond to any real causal processes.

<sup>vi</sup> Although Newton’s theory of gravity had always lacked a mechanism – even he was concerned that it required instantaneous action at a distance, which is implausible. It was therefore an incomplete theory. During the twentieth century, Einstein’s theory of general relativity provided a solution for this problem. It is therefore inaccurate, as is often casually done, to refer to Einstein as having “overthrown” Newton’s theory that was previously thought to be a complete account – there was always a severe problem with it. The equations worked well for most purposes though, and still do.