Is Economics a Science? Well, Not Yet.*

Ricardo Dahis

Northwestern University

February 10th, 2018

Abstract

Is economics a science? Answering this question is not only necessary for philosophical clarity, but also crucial for knowing how seriously to take economists' claims and advice about public policy. Nevertheless, even among practitioners and academics, consensus is nonexistent. This paper resolves the conundrum in two steps. First I discuss some epistemology of science, defining clearly various concepts necessary to the debate. Several fallacies are clarified, such as "a theory may never be proved true, but only not falsified", "a model is useful because it simplifies reality" or "data mining is bad". In light of solid philosophical ground, I then discuss the practice and methodology of modern economics. The answer to the title question is a perhaps disappointing, but realistic, *not yet*. I conclude with prescriptions for a path towards a more scientific discipline.

*I would like to thank Flavio Abdenur, Kym Ardison, Mariana Carvalho, João Machado, João Thereze Ferreira and Paulo Orenstein for helpful comments and suggestions. Correspondence: rdahis@u.northwestern.edu.

There is a story that has been going around about a physicist, a chemist, and an economist who were stranded on a desert island with no implements and a can of food. The physicist and the chemist each devised an ingenious mechanism for getting the can open; the economist merely said, "Assume we have a can opener"!

A policeman sees a drunk man searching for something under a streetlight and asks what the drunk has lost. He says he lost his keys. The policeman asks if he is sure he lost them here, and the drunk replies, no, and that he lost them in the park. The policeman asks why he is searching here, and the drunk replies, "this is where the light is".

At dinner with the queen of Sweden [. . .] she asked me what I won the Nobel Prize for. When I said "Quantum Physics", she said, "Oh, we can't talk about that, because nobody understands it." and I said "On the contrary, we know quite a lot about quantum physics, and THAT'S why we can't talk about it. It's everything else we DON'T know about - like how to solve poverty, and lower crime, and stop drugs, that we CAN talk about!"
— Richard Feynman, after receiving the Nobel Prize

The question of whether or not economics is a science is decades-old, but its answer still seems elusive to most people. And the question is consequential: how much should we trust our economic predictions and explanations? Put differently, do we trust our understanding of social phenomena as much as we do with physics or chemistry? Are we willing to use economics' conclusions to guide public policy, both locally and worldwide?

On the one hand there are reasons for a pessimistic answer. Economists have a hard time making good predictions, be it next year's inflation or the Great Recession in 2008. The classical saying goes that with two economists one will always find three opinions. Economists may sometimes agree on solutions to basic policy issues, but we're very far from consensus about the more complicated ones. On the other hand there are reasons for optimism. Economists use the vocabulary of science: words like theory, data and experiments are used in everyday discourse and practice. Economic theories are full of math and rigor, and genuine intellectual effort is put into distinguishing what links between models and data can be inferred.

Therefore it's not a surprise that the question still seems to generate controversy. In particular, star economists have recently expressed their views online but haven't reached a conclusive and satisfying answer (see Chetty (2013) and answers by Krugman (2013) and Wang (2013), or Shiller (2013)). In another manifestation causing stir, Romer (2015) criticized the problem of "mathiness" in economics.[1]. In a series of posts on economic methodology, Noah Smith also discusses many of the lazy[2], good and

[1]See Smith (2015c).
[2]See Smith (2015d).

bad[3] criticisms to economics. There is still misunderstanding about the potentials and limitations of experiments in economics (Deaton and Cartwright, 2017), and about the interplay of different methodologies such as "structural" vs. "atheoretic" (Keane, 2010).

But what is this discussion missing? Why does there seem to be a philosophical confusion among economists? In this essay I use some epistemology of science with precise definitions, along with examples of how modern economics is made, to resolve this conundrum. The answer to whether economics is a science or not, as the title of the essay warns, is a perhaps disappointing, but realistic, *not yet*. Nevertheless, I argue below that the field is doing progress and that it sometimes already reaches a weaker status of science.

The rest of the essay is organized as follows. In section I I develop some simple philosophy of science to clear various confusions present in texts and discussions about the topic, and provide a definition of science. In section II I dive deeper into how modern economics is made and discuss whether it is a science or not. Section III concludes with some final comments and prescriptions.

I Some epistemology and philosophy of science

If the goal is to classify economics into some category, we better understand what that category is. And to properly discuss what science is, we need a basic epistemological framework set in place.[4] So let's start with some definitions.

I.A The basics

The first is *truth*, which in philosophy also requires a definition. And an appealing one is *truth as correspondence* (Marian, 2016).

Definition 1. A claim or statement (with propositional value) is *true* if it corresponds to reality.

Not all statements have propositional value, e.g. "what a beautiful bird". Reality may be measured through observations, experiments, etc., and is loosely understood to be the objective source of stimuli perceived through human sensibility. Reality contains the sun, the trees, the buildings made of concrete that you and I can touch.[5] Not all claims are correspondent with reality. For instance, nowadays we know the claim "the Sun rotates around the Earth" to be false. Other claims we may not be so sure, e.g. "eggs are good for your health". General claims may be false, but weaker (more

[3] See Smith (2015b).

[4] A thorough presentation of epistemology may be found in Goldman (1988).

[5] Notice the implicit assumption of philosophical *realism*, along with other assumptions in the *default* position (Searle, 1999).

specific) versions of themselves may be true. Moreover, different types of claims require different types of evidence to prove them true. Consider the following sequence of claims.

1. All humans love ice cream.

2. All economists love ice cream.

3. Ricardo Dahis loves ice cream.

Proving the first true requires asking every single human being who ever lived. Proving the second requires asking all economists who ever lived. Proving the third is easy.[6]

Claims in science are often expressed in and/or derived from theories or models. Nevertheless, much confusion revolves around what are both these objects, and their relationship to truth claims. The following definition is natural.

Definition 2. A *theory* is a set of claims with propositional value.

A theory's claims are often named *hypotheses* or *assumptions*. These hypotheses may be explicit or implicit. They may even include causation claims. Besides, a theory usually contains claims about *entities*, such as the Sun, a person, or an atom. Entities are also set of descriptive claims, which can often be made implicit when stating theories – e.g. a person has one mouth and two eyes. Theories imply *conclusions*, or *predictions*, which are also claims, deducted from the hypotheses by logical rules of inference.

The definition above requires several more comments. First, notice that theories don't necessarily have to concern real or empirical entities. We may make abstract hypotheses about imaginary beings in Mars and use logic to conclude something about their imaginary lifestyle. This definition also classifies mathematical theorems as theories. Second, the same prediction may be reached by a different set of hypotheses. For example, we can conclude that an airplane can fly because of complicated engineering or because it is being controlled by invisible flight gods. We could also deduce that climate change is happening either because of human activity or natural changes in the Earth's biosphere. We'll discuss how to evaluate hypotheses below.

Moreover, notice that we have restricted a theory to include deductions only from valid rules of logical inference. Some examples of such rules are *reductio ad absurdum*, *modus ponens*, *modus tollens*, etc. If, from a hypothesis about Brad Pitt's height, we then conclude that the giraffes are insects, then, well, not only we have lost our minds, but our theory also doesn't make any sense. Fourth, sometimes theories have implicit assumptions such that, without those, the conclusions don't necessarily follow, and nevertheless such assumptions are not clearly stated. Lastly, the definition above is

[6] It is true.

very broad. Almost any propositional claim we can think of can likely be expressed as a set of assumptions. And logical deductions may follow.[7]

Definition 3. A *model* is a theory whose hypotheses contain some type of simplification in comparison to the real world.

Therefore every model is a theory, but not necessarily every theory is a model - some theories involve no simplification whatsoever. On the other hand, to study climate change one needs to simplify how weather systems work in order to get any reasonable predictions out of the model. Nevertheless, it's not rare to see people using the terms *theory* and *model* interchangeably, and that creates confusion.

When discussing the merits of different theories in science, several classification dimensions are employed. Let's discuss each one of them because they are not necessarily related to one another. The most important ones for the purposes of this paper are verifiability and falsifiability.

Definition 4. A theory is *verifiable*, or *provable*, if, at least in principle, its hypotheses may be directly compared with reality.

Notice then that a theory is proved true if all its hypotheses are proved true. And it is proved false if at least one of its hypotheses is proved false. Moreover, a theory may be unverifiable for two reasons. Either (1) one cannot, in principle, measure what the assumptions posit, or (2) one can, in principle, measure what the assumptions posit but it's unfeasible to do so. For example, proving a theory about stars would require observation of every single star that ever existed and that will ever exist in the whole universe. Or, as Hume (1738) famously discussed as the problem of induction, one cannot infer that "instances of which we have had no experience resemble those of which we have had experience". It's not because we observe the Sun rising every morning that we can prove it will rise tomorrow.

In addition, not all theories are verifiable. Yet, unverifiable theories may generate predictions we would like to take seriously and potentially call scientific. Which brings us to our next definition.

Definition 5. A theory is *falsifiable*, or *testable*, if, at least in principle, its conclusions may be directly compared with reality.

This is a restatement of what was first formalized by Popper (1959), in its beautiful solution to the induction and demarcation problems. Popper separated science

[7]The word *theory* holds a different status in philosophy of science as compared to common parlance. In colloquial language a theory is something yet unproven, being purely hypothetical and not to be taken as fact. Therefore, even theories that are repeatedly shown to make precise predictions, such as the theory of Evolution by Natural Selection or Einstein's Theory of General Relativity, which most scientists would take as fact, are dismissed as "only a theory" (most notably within religious groups). But, in science, a theory may be verified true, or not verified false. More on this point below.

from nonscience, pseudo-science and other products of human activity (such as art or beliefs) by requiring a scientific claim to be "capable of conflicting with possible, or conceivable observations".

Let's further parse this definition with care. First, a theory may be refuted in two different ways: either it is verified false, or it is falsified. A theory is falsified if its predictions conflict with observation or experiment. Being falsified implies at least one of two properties: (1) if all the theory's assumptions are verified true, then the theory is *incomplete*, or (2) if the theory contains unverified assumptions, then either at least one of these is false, or the theory is incomplete.

Second, we can use these two definitions above to clarify a common fallacy in science, namely that "a theory may never be proven true, but only not falsified". Some theories may be verified true by straightforward observation and measurement. For example, today we know the claim "water is made of two parts oxygen and one part carbon" to be true. Or "the Earth rotates around the Sun".

Nevertheless, for unverifiable theories, as Richard Feynman elegantly described it, scientists have degrees of certainty and possible beliefs over their [unverifiable] theories. As we repeatedly test the theory, across different contexts and moments in spacetime, we can grow more confident that the theory is true. Besides, the sole judge of whether a theory is any good or not is nature. It doesn't matter what your name is, how famous or rich you are: if the theory's predictions disagree with experiments and observation, it is wrong (Feynman, 1964).

Third, science can only answer the *how* questions, those of cause and effect. Questions such as "how does Earth's gravity affect the Moon's orbit?" are answerable by science. But the big *why* questions, in the sense of purpose as understood by our human intuitions, are not in the realm of scientific activity. All physicists nowadays understand the concept of inertia and how it works, but there's no explanation for why inertia exists in the first place.

Finally, the elaboration of new theories may happen in different ways. A common one is logical abduction: finding a hypothesis that implies an empirically observed phenomenon ex ante.[8][9] As new and possibly contradictory evidence is collected from observation and experiments, theories may be revised to fit the new facts, or be altogether discarded in favor of alternative ones. If the accumulation of "anomalies" is big enough, it may even cause a paradigm shift in the field (Kuhn, 1962).

Being testable is what defines a theory as being scientific, but is not the only quality by which to judge it. In fact, the other various classifications a theory may hold are

[8] Alternatives are induction, deduction, or simply guessing.

[9] A common question in science is "what theory can explain the observations?". To be clear, a theory *explains* an observation if it generates predictions corresponding to the observation. Finding a theory that explains some observation is, therefore, abducting a set of assumptions that accounts for the empirical observations.

a source of confusion in economics, where practitioners confuse expressions like predictive power or realism with testability. Let's define the most common dimensions people usually care about. Notice that they may not always intersect one another.

Definition 6. A theory has *predictive power* if its predictions correspond to reality.

This is a very desirable feature for a theory to have. After all, we like theories that can reliably predict events in the real world. Quite naturally there is heterogeneity across theories, with some having more predictive power than others. And it's important to point out that a theory may seem completely absurd and yet have good predictive power. A classic example explored by Feyerabend (1975) is geocentrism, the cosmological theory that the other planets and Sun revolve around the Earth. Before Heliocentrism was advanced by Copernicus and Galileo, the former theory still made reasonable predictions about Sun's movement relative to the Earth.

We can now clarify another common fallacy, namely that reverse-engineering theories to fit empirical observations, or doing "data-mining", is always bad practice. Abducting theories from observation is precisely what scientists did for centuries when crafting theories. That is not the problem. The real issue is that in some fields it is difficult to replicate experiments and, thus, difficult to do out-of-sample testing. One cannot then differentiate curve-fitting from a predictive theory that would work in more contexts. But fitting theory to data is legitimate when one can iterate multiple times in new experiments.

Definition 7. A theory is *realistic* if its assumptions correspond to reality.

Theories about unrealistic entities are therefore unrealistic. A theory only involving observable aspects of people is realistic, while a claim about elves is not. Theories may also have different degrees of realism, and may include at the same time both realistic and unrealistic assumptions in its structure. Additionally, a theory's realism can be constrained by the observational capacity of its day. We nowadays consider Darwinian Evolution a realistic theory, but it may come the day that we discover new aspects of biology that will render the theory less so.

This is a good moment to issue a correction to a common fallacy in science. A model is often described as a map, and its "usefulness" is measured by its ability to simplify reality and expose the main forces behind a phenomenon. The argument states that a map which contains every aspect of reality is useless, since the map is then basically reality itself all over again. And a model that tries to capture all the complexities of reality will neither yield sharp predictions nor serve to isolate smaller mechanisms in a bigger context and aid understanding. Yet, the implication that models with simplifications are "better" than more complicated ones is a fallacy.

Simplification, normally done by making stronger assumptions, may help humans understand complex phenomena by reducing them to a set of simpler and smaller mechanisms, possibly increasing the model's intelligibility and elegance. It may also indeed increase the model's predictive power, but this increase is not a necessary condition. Evidently, simpler models (or theories) are not any more scientific than more complex ones. Simplification may be a useful step towards building scientific and predictive theories, but it's not more than that.

Let's proceed with mathematization.

Definition 8. A theory is *mathematized* if its assumptions are expressed in the language of mathematics.

As Wigner (1960) famously put it, mathematics possesses an "unreasonable effectiveness" in the sciences. Math not only is an instrument for clarity and logic, but in fact its concepts have applicability far beyond the context in which they were originally developed. Not surprisingly mathematics is used extensively in every field of science, with some theories naturally being more mathematized than others. While some theories may be stated in a simple logical $P \implies Q$, others such as string theory require complicated math that may even get you a Fields medal for it.[10]

Nevertheless, there is no intrinsic quality in a theory being mathematized. The value of mathematics for science is in its usefulness, not in its pure existence. Besides, the use of mathematics doesn't imply that a theory is scientific, much less *good* in any sense. Math is simply the language through which scientists very often express hypotheses, deductions and conclusions.

Definition 9. A theory is *beautiful*, *elegant*, *intelligible* or *internally consistent* if humans say so.

Jokes aside, these are four dimensions that theories are often evaluated by and which have no objective measure. People frequently classify theories as beautiful based on some aesthetic sense of purity, austerity and symmetry. This trait is notably present among mathematicians, as expressed by Bertrand Russell and Paul Erdos. Likewise, theories are appreciated by their elegance. This concept is harder to express, but is related to a sense of explanatory simplicity in the face of the complex phenomena studied. The Euler formula $e^{i\pi} = -1$ is seen among many as a supreme form of beauty and elegance in mathematics. Among biologists the theory of Evolution is also judged elegant, given its enormous explanatory power and simple mechanistic theoretical structure.

Additionally, theories (and the universe) may be intelligible for human intelligence and intuitions. And this is sometimes surprising, as Einstein expressed by saying that "the most incomprehensible thing about the universe is that it is comprehensible". Interestingly, despite yielding predictions accurate to a dozen decimals, physicists are

[10] By the definition above, nowadays string theory cannot be classified as scientific, given its lack of testable predictions in a three-dimensional world.

still trying to make sense of quantum physics, where very small particles show entanglement, non-locality, probabilistic positions in space, etc. Finally, theories may or may not be internally consistent, which loosely means it having the same language and concepts to describe and predict a wide variety of phenomena. All four dimensions above are merely human subjective evaluations, which are unrelated to a theory being scientific or having good predictive power.[11]

Finally, with these basic definitions from philosophy of science in place, we are ready to deliberate over and define science.

I.B What is science?

The word *science* itself is originated from *scientia* in Latin, which means "to know", according to the Oxford English Dictionary. This notion encompasses all branches of knowledge, and would for example count art, or inner knowledge, as science. Clearly this is too broad for modern standards, so it cannot serve our purpose in this essay.

A second possible definition is science as a "systematically organized body of knowledge on a particular subject". This definition is the one commonly used in epistemological coffee-breaks, or in informal descriptions of disciplines. Nevertheless, it will also not do. Simply having a systematically organized body of knowledge on a given subject doesn't differentiate the methodologies through which such knowledge is acquired. According to this definition, fields like astrology, which have long histories in human intellectual endeavors and plenty of books written on the subject, could be described as science. However, the methods used to accept a claim as valid matter: we need to differentiate truth claims based, for instance, on introspection, on observation, or experimentation. This definition is also the source of the deep confusion caused by the term *social sciences*, which immediately conflates the study of social phenomena with science. We should replace this expression, instead, with *social studies*.[12]

A third definition, which is still a weak definition for science, contains the ones above and makes a fundamental restriction.

Definition 10. *Science* is a set of testable theories, a set of observations and evidence supporting these theories, and a set of systematic practices and protocols to generate measurements, categorize data, formulate new theories, and perform experiments.

[11] Although it is widely recognized, and culturally institutionalized with Occam's razor, that, given the same predictive power, simpler theories should not only be preferred to more complicated ones, but also tend to be more likely true.

[12] *Social science* is a misleading expression and should therefore be substituted to something like *social studies*. The expression implies science as merely an activity of systematic investigation, and doesn't clarify the requirement of contrasting falsifiable theories with data. Since not all fields within the social studies fit this criterion, we risk creating a consequential confusion. Notice additionally that, for example, physics is not called physical science. The word science should be used only as a qualifier to fields. The same argument applies to political science. The field should be called *politics* and its practitioners *politists* (in the same spirit as economics and economists or physics and physicists).

The crucial restriction is that scientific knowledge is obtained through testable explanations that generate predictions about different phenomena, be it natural or social. The method used to test explanations is the *scientific method*, which is composed of a set of elements: observation, measurement, experimentation, data analysis, replication, peer review, among others. Fundamentally, science is a search for invariance in nature. And this invariance can be described as abstract *laws* (expressed in logical and mathematical language) that are then used to generate testable predictions about empirical phenomena.

Additionally, the definition above is weak because it is still too inclusive. It requires a body of knowledge to use the scientific method and to structure its claims as predictions about empirical phenomena, but it imposes no measure of efficacy or predictive success over the discipline's claims. In other words, it would classify as science a field that "does everything right", that searches for invariance across spacetime, but that doesn't reach conclusive laws.[13][14] An example of (as of today) unsuccessful science is the study of nutrition, from which every other day the recommendation about eating or avoiding eggs and chocolate changes.

The discussion above motivates our final definition of this section.

Definition 11. A science is *predictive*, or *successful*, if its theories have predictive power.

II Is economics a science?

Now let's move to the substantive question posed in this essay's title. In the interest of clarity, let's start once again by trying to define economics.

II.A What is modern economics?

The word *economics* originates from the greek *oikonomika*, meaning the "study of household management" (Leshem, 2016). Obviously the meaning of the word has changed in important ways, so let's use a modified version of the definition given by the American Economic Association.[15]

[13] The lack of success may result from different causes, and we'll discuss some issues specific to economics below. Sometimes the phenomena studied have no invariant expressions, or its patterns change too much across time and" space.

[14] Feynman (1974) discusses the differences between science and pseudoscience and classifies a subset of weakly scientific fields as *cargo cult science*. These enterprises fail to be scientific because of many reasons, one being a lack of investigative integrity. The idea, when designing theories and experiments, is to try to give all of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another. Impediments to integrity, such as authority in publication decisions or lack of replicability, slow down a successful search for empirical laws.

[15] See https://www.aeaweb.org/resources/students/what-is-economics.

Definition 12. *Economics* is the study of scarcity, the study of how people use resources, or the study of decision-making. Economics is also the study of incentives (be it monetary or not) and how people react to them.

Economics often involves topics like wealth, finance, recessions, and banking, leading to the misconception that economics is all about money and the stock market. Actually, it's a much broader discipline that helps us understand historical trends, interpret today's headlines, and make predictions for coming decades. The scope of the discipline was immensely enlarged in recent decades, with economists now asking questions about politics, marriage, culture, kidney exchange, law, among others. In fact the methods of modern economics apply to any context where there are costs and benefits and a decision has to be (optimally) made.

The objects of study of economics are mainly humans and their social organizations. And studying humans is, in a way, fundamentally harder than studying physical phenomena. Social phenomena are modulated by hundreds of variables and billions of individuals that interact in complex, dynamic and often unpredictable ways. Nevertheless the problem is not complexity per se[16], but isolating it in search of invariance across time and space. A conclusion that a certain cultural trait causes people to build their cities in a certain way in Germany during the 19th century hardly ever serves as a predictive theory in other contexts. And this difficulty has a cause. It's frequently hard, expensive and/or unethical for economists to come up with controlled experimental data to test their theories. History unfolds itself in front of us and we can rarely experiment with it. It's no surprise then that *ceteris paribus* analyses and experimental replication are virtually inexistent in the discipline.[17]

The practical result is that the data used in economics is often observational or choice data, in which agents have already made decisions and one only observes the result of a possibly complicated causal process underneath.[18] The challenge, then, embodied in the field of econometrics, is to draw valid inferences from non-experimental data, where a dozen variables may be affecting the one the economist wants to study.

In sum, testing theories in this context is hard. To do progress, economists had to get creative. And among the great minds in economics there was Paul Samuelson. Along with his many contributions to the field, Samuelson (1938) pioneered the idea of *revealed preference*, which goes like this: let there be two bundles of goods, a and b, available in a budget set B. If it is observed that a is chosen over b, we say that a is

[16]Branches of science and math, such as Chaos Theory or Dynamic Complex Systems, study immensely complex objects and are still able to make predictions about their properties. In studying the behavior of a gas spreading inside a room, one usually cannot say anything about the position of a single particle (given the randomness involved in the particles' interactions) but can still predict the density at each physical position after a certain time.

[17]Exceptions exist in fields that use controlled lab experiments, where replication is more feasible.

[18]Consequently, virtually any empirical test designed to refute economic theories has to involve data the theories were designed to explain, i.e. choice data.

(directly) revealed preferred to b. This, together with axioms of transitivity and completeness, allowed economists to abstract away from any preference formation underlying process and build a decision theory requiring agents only consistency between choices. Now economists could interpret choice data as originated from a principled revelation of preferences and, with that, extrapolate how agents would choose in different environments (Manski, 2007).

Some years later, Milton Friedman would also have a profound impact on the profession. In two very influential essays that remain controversial to this day, Friedman and Savage (1948) and Friedman (1953) discussed the methodology of positive economics. In it Friedman discussed many important topics that I won't cover here, such as the distinctions between normative and positive economics, but, in particular, he and Savage set the stage for a long-lasting confusion among economists thinking about epistemology. The parts I'm referring to are the "judge a theory by its predictions" and the "as if" arguments.

In Friedman (1953) the author grants that "a theory cannot be tested by the realism of its assumptions" (p.16), and goes on to conclude that "the ultimate goal of a positive science is the development of theory or "hypothesis" that yields valid and meaningful (i.e., not truistic) predictions about phenomena not yet observed." (p.5).

And in Friedman and Savage (1948), in arguing that economic models of maximization behavior are successful *as if* approximations of decision processes, the authors offer the famous billiard player metaphor. The full argument from p.298 is quoted below, given its historical importance.

> The hypothesis does not assert that individuals explicitly or consciously calculate and compare expected utilities. Indeed, it is not at all clear what such an assertion would mean or how it could be tested. The hypothesis asserts rather that, in making a particular class of decisions, individuals behave as if they calculated and compared expected utility and as if they knew the odds. The validity of this assertion does not depend on whether individuals know the precise odds, much less on whether they say that they calculate and compare expected utilities or think that they do, or whether it appears to others that they do, or whether psychologists can uncover any evidence that they do, but solely on whether it yields sufficiently accurate predictions about the class of decisions with which the hypothesis deals. Stated differently, the test by results is the only possible method of determining whether the as if statement is or is not a sufficiently good approximation to reality for the purpose at hand.
>
> A simple example may help to clarify the point at issue. Consider the problem of predicting, before each shot, the direction of travel of a billiard ball

> hit by an expert billiard player. It would be possible to construct one or more mathematical formulas that would give the directions of travel that would score points and, among these, would indicate the one (or more) that would leave the balls in the best positions. The formulas might, of course, be extremely complicated, since they would necessarily take account of the location of the balls in relation to one another and to the cushions and of the complicated phenomena introduced by "english". Nonetheless, it seems not at all unreasonable that excellent predictions would be yielded by the hypothesis that the billiard player made his shots as if he knew the formulas, could estimate accurately by eye the angles, etc., describing the location of the balls, could make lightning calculations from the formulas, and could then make the ball travel in the direction indicated by the formulas. It would in no way disprove or contradict the hypothesis, or weaken our confidence in it, if it should turn out that the billiard player had never studied any branch of mathematics and was utterly incapable of making the necessary calculations: unless he was capable in some way of reaching approximately the same result as that obtained from the formulas, he would not in fact be likely to be an expert billiard player.
>
> The same considerations are relevant to our utility hypothesis. Whatever the psychological mechanism whereby individuals make choices, these choices appear to display some consistency, which can apparently be described by our utility hypothesis. This hypothesis enables predictions to be made about phenomena on which there is not yet reliable evidence. The hypothesis cannot be declared invalid for a particular class of behavior until a prediction about that class proves false. No other test of its validity is decisive.

In sum, Friedman and Savage argued that economists could use the utility hypothesis without necessarily checking if that's what is really inside the billiard player's brain. Additionally, the theory could then be judged solely on its predictive power. And, in terms of predicting the shot's outcome, this argument sounds reasonable. There would be no loss in assuming the player (or agent) behaves rationally, i.e. as if maximizing some objective function, if we only care about prediction.[19] So why is this construction epistemologically confusing for judging whether or not economics is a science?

[19] Why do economists often assume rationality in their models? This question deserves more than a footnote, but some of the reasons are: (1) it makes models easier to solve mathematically, (2) it allows for separating institutional/market structures from choice features, (3) the Ana Karenina's corollary (there is one way to be rational, but many ways not to be), and (4) it is a good approximation for decision making in certain contexts.

The point is that, in science, if a theory is tested and not refuted, we grow more confident of it being true (in the correspondence sense). However, *as if* assumptions are often unverifiable and unrealistic. A non-refuting test to, say, utility theory does not give us any more belief that the theory is true, given the unrealism of its hypotheses. Such a test would make us grow more confident with the theory's predictive power, but not its descriptive accuracy of the real world.[20] When only one theory is available, we take what we have. But if two theories have the same predictive power, then an analogous Occam's Razor should apply and we should prefer more realistic theories.

In addition to a philosophical confusion, the methodological consequences of Samuelson's and Friedman's contributions were vast. The tone was set for the type of modeling economists would do in the future, and many components of science in section I would be reinterpreted. For instance, economists use a weaker concept of prediction than most scientists would be satisfied with. Economists most often only make qualitative predictions such as an effect's sign, and use vague language like "short-term" or "long-term". And weak prediction requirements imply empirical tests with little bite. This in turn drastically reduces the probability of theories being refuted and thus refined and improved.

A good example is the *as if* argument itself. One could do away with the deeper descriptive question about whether people actually behave rationally and rather ask the more pragmatic one, about whether the rational *as if* model can predict behavior well or not (Chetty, 2015; Fama and Thaler, 2016). And, empirically speaking, the *as if* argument fails to hold in contexts where agents have no training, when there's little repetition or immediate feedback to mistakes. Nevertheless the profession's reaction speed in formulating new theories was slow.[21]

Moreover, this methodological agenda propelled economists into a culture with too much acceptance of strong and unrealistic assumptions (e.g. the can opener), where internal consistency, simplicity and elegance as values prevail over real predictive power.[22]

[20] This is not to understate the power and usefulness of *as if* arguments in various fields of science, in particular in situations where evaluating an assumption's realism is tricky. For example, in physics one observes Newton's apple falling and explains it by saying that it falls *as if* a force was pulling it down. The concept of force itself is unintuitive and "unrealistic", although it perfectly serves the purpose of predicting the apple's trajectory in space. Other examples are General Relativity, in which light travels through space *as if* this space was curved by gravity. Or in biology where genes behave *as if* they were maximizing their inclusive fitness. Or finally again in physics where light travels *as if* minimizing the time to travel between two points.

[21] This culture has been changing in the past two decades with the growth of behavioral economics and access to larger and better datasets. For interesting recent methodological discussions, see Thaler (2016), Caplin and Schotter (2008) and Glimcher and Fehr (2014).

[22] Once a tribe that had similar tastes and rituals to the economics profession was studied. The ethnographic analysis is fascinating, and can be found at Leijonhufvud (1973).

II.B Fables, analogies, chameleons, and more

Various authors have noted that theories and models in economics are used and interpreted not in accordance with standard practice in the natural sciences or psychology. In this section I discuss some of the more recent views about theory that circulate among and represent broad views of academics and practitioners. Most of them are not particularly helpful, and may actually be harmful by causing confusion over the nature of theory as discussed in section I.[23]

Rubinstein (2012) describes economic models as *fables*. Economists create worlds between reality and fantasy, which "can be free from irrelevant details and unnecessary diversions", and thus "help us learn a lesson from the story". Moreover, "models can be denounced for being simplistic and unrealistic, but modeling is essential because it is the only method we have of clarifying concepts, evaluating assumptions, verifying conclusions and acquiring insights that will serve us when we return from the model to real life". Rubinstein recognizes that using formal language "creates the illusion of being scientific", mainly by "conceal[ing] from the layman the assumptions the model uses".

In an earlier publication, Rubinstein (1998) offers four views of economic models: (1) "models are aimed to predict behavior", (2) "models are normative, in the sense that they are supposed to guide the economist in giving advice to economic agents about what to do", (3) "models are exercises intended to sharpen economists' intuitions when dealing with complicated situations", and (4) "models are meant to establish "linkages" between the concepts and statements that appear in our daily thinking on economic situations". Each of these views has a ring of truth to it, but none address the status of economics as a science or not.

Gilboa et al. (2014) view economic models as *analogies*. In particular, they first distinguish reasoning as "rule-based" and "case-based", and then "suggest that economic reasoning is partly case-based, and that one role of theory is to enrich the set of cases". Case-based reasoning would only require a judgement of similarity between the past cases and the present one in analysis (provided a similarity distance function). And models would serve as analogies, i.e. theoretical cases which economists can draw from to compare to new situations and make predictions. Although interesting, the distinction between cases and rules is, in fact, an illusion. Descriptions of rules and cases are both theories as sets of hypotheses that generate predictions. Cases are a rule, but with more, or different, assumptions (often unknown or not stated by the theorist) as compared to more general rules. Testing a theory, be it a case or a rule, always requires judging similarity of its assumptions to the empirical setting. Having

[23]For a more extensive literature review on economists' and philosophers' views on the use of theory in economics, see Gilboa et al. (2014).

theoretical cases to draw from helps making economics more of a science only as far as it aids formulating testable theories. And testability is what really matters.

Finally, Pfleiderer (2014) views economic models as potentially *chameleons*. Chameleon models are those put forth as "saying something about the real world" but, when criticized, are defended as being just thought experiments – theories that shouldn't be judged by the realism of their assumptions. Theories become chameleons by avoiding the proper empirical tests, or "filters", and by being taken seriously "until proved guilty". As the author recognizes, what really determines the scientific value of a theory is being compared against reality and passing or not the test. The chameleon definition is useful for identifying ill-advised policy recommendations.

Among the many other views of economic theory not expressed here, one is particularly confusing. If all else fails, economic theories can always be interpreted as *normative*, a view that puts them in different arenas of validity criteria.

II.C So, is it a science or not?

Despite all critical points made above, there is space for good news. In recent years economics has become more and more data-driven, as a byproduct of the surge of new large and better datasets and exponentially growing and cheap computational power. Nowadays economists run thousands of regressions and statistical analyses in minutes, and can simulate significantly complicated models in computer clusters in the cloud: a feat our predecessors would hardly even dream about. The field is becoming increasingly aware of the importance and challenge of testing theories and refuting them when necessary. In fact the number of experiments performed and empirical papers published has never been higher.[24]

And, building on the seminal paper by Rubin (1974), important innovations in the study of causal inference were incorporated into the discipline. Methods such as Regression Discontinuity Design, Difference-in-Differences and Instrumental Variables became the workhorse for reduced-form applied research, together with other methods in the Industrial Organization literature. The critical literature on (partial) identification and specification-testing has flourished in recent decades (Manski, 1995; Leamer, 1983). Natural experiments are also increasingly popular since they induce exogenous variation in real-world phenomena (and not inly in the lab). In fact it has been recently argued, however controversially[25], that economics is going through a "credibility revolution" (Angrist and Pischke, 2010). And, not surprisingly, a number of recent initiatives such as the J-PAL and IPA are pushing the agenda for evidence-based policy interventions in countries worldwide. Registration of pre-analysis of ex-

[24]See Smith (2015a).

[25]See the discussion in Deaton (2009), Heckman and Urzúa (2010) and Imbens (2010).

periments to avoid data-mining and publication bias is quickly becoming the norm (Christensen and Miguel, 2016). Journals with space for replication studies is not yet a reality but are already discussed in the profession.[26]

Nevertheless, economics is a bundle of different fields, each having its own culture, standards and goals. For example, modern macroeconomics uses modeling choices such as infinite-horizon dynamic optimization, rational expectations, overlapping generations, financial frictions, Kalman filters, among others. The standard methodology is to write a model of the economy in which technology, preferences and random shocks' distributions are specified. One then provides an equilibrium definition and solves the model. The final steps are then parameter calibration, matching data distribution's moments and simulating impulse response functions. However each model has its own setup, and attempting to include more than a few real-world complexities to the model quickly makes it analytically or computationally intractable. Given the lack of controlled experimentation and of good data on macro phenomena, it seems unlikely that the field will reach the required standards of definitions 10 or 11.[27]

The answer to whether or not economics is a science is, therefore, not simple. The discipline can be considered a science only as far as economists craft testable theories, compare its conclusions to reality and refute the ones not passing the test. And this quality is very context-dependent, with gradations across and within fields. There's also heterogeneity in the other less consequential dimensions from section I: some parts of economics are more mathematized (e.g. choice theory), while others may be more realistic (e.g. policy evaluation). So the answer to economics being a science or not is a humbling, but solid, sometimes.

Now, is economics a predictive science? Well, not yet. The field as a whole is far from having the descriptive and predictive success of physics or molecular biology, and it's not clear now if such progress will ever happen. And, despite the advances discussed previously, we can list several reasons for this state of affairs, taking advantage of all definitions and discussion developed above.

First, economics suffers from many of same difficulties other fields of science also have.[28] Too many studies are poorly designed, replication results are rare, peer review is not double-blind and is very often biased, etc. Publication and journal politics can stop publication of critical works, and few journals require publishing the paper's data and (readable and replicable) code.

More specific to economics, Chang and Li (2015) attempt to replicate 67 papers

[26]Empirical work has two main purposes. One is to cumulate evidence and data that eventually informs generating theories. The other is to properly test theories. But the goal of science is to craft predictive theories that are as simple as possible, and not to endlessly test and re-test theories. There is no point spending resources in testing for the Earth's gravitational pull once we know what it is.

[27]Although identification tides in macroeconomics are changing. See Nakamura and Steinsson (2018).

[28]See http://www.vox.com/2016/7/14/12016710/science-challeges-research-funding-peer-review-process.

in well-regarded economics journals and succeed in replicating the main qualitative results for only 22 (33%) without contacting the authors. With the help of authors this number increases to 29 out of 59 (excluding papers with no data or software). In other words, this sample of papers (which are likely representative of the field) is not widely replicable.

But the strongest reason, as discussed in the beginning of this section, is that the object of study in economics may not be truly amenable to the scientific method. Even if we did "everything right", it would probably still be hard to find general laws that applied to the plethora of phenomena economists study. Additionally, at least, economists could incorporate more of the culture in disciplines such as neuroscience or psychology. Both also study human beings – including behavior and decisions – but start from "simpler" questions, such as how vision works or how to build a computer that plays poker well. By isolating smaller pieces of complexity at a time, these fields are making progress and are progressively becoming predictive sciences. And, as the scientific knowledge grows, one can start asking bigger and more ambitious questions, such as how do groups of people behave in market settings or how institutions modulate economic growth. Economists want to sidestep the science-making: they ask the most interesting questions but often end up with only frustrating answers.

III Final comments

Recent decades and the future point to a convergence of the social studies. The distinctions between politics, economics, anthropology or sociology are becoming increasingly a matter of question type and emphasis instead of the methods used. The use of statistics, causal inference, etc., will more and more be part of the vocabulary of any social investigator. And, presently, economics is arguably the methodologically most advanced of the social studies. The level of rigor and effort put into precise definitions, identification of parameters, construction of clear links between theoretical constructs and data, among others, indicates that the profession is serious about trying to become a science. We should not mistakenly judge the contemporaneous lack of success as incompetence or sloth across the field, quite the contrary. We have already argued in section II that the object of study tackled by the social studies is complex, dynamic and with little invariance across space and time, which makes the investigator's task difficult.

It's hard to predict today if the discipline will ever achieve the status of a predictive science. The nature of the object of study will not change, but maybe methodological innovations allied with more data and computational power will serve as keys to progress. Nevertheless, however hard we push the boundaries, for the time being we will have to accept that economics' claims are too often still educated guesswork and smart stories; and that the policy advice economists produce does not have the authority of a physicist computing the force of Earth's gravity. Nevertheless, something is better than nothing. Whatever the science status of economics, at the end of the day someone has to choose the country's interest rate or fiscal policy. And economists are the most qualified to advise such decisions.

The way forward to an economic science will involve a list of steps. Economists have to be even more willing to test their theories and refute them when necessary. Qualities like internal consistency, mathematization and elegance should have a second-order priority in the search for new theories. The assumptions made in models should become progressively more realistic, so that they can be directly applied in empirical contexts that fit their framework: the *as if* argument should stop being a refugee for unrealistic theory. And (true) prediction should be our first goal: understanding and distinguishing the mechanisms should only then become the priority.[29] If economics is ever to become a predictive science, the transformation will happen only if we push our methodology in the right directions.

[29]Evidently these two activities mutually influence and advance one another. An interesting discussion on how machine learning and causal inference relate, and how the first can bring economics to new predictive levels, can be found at Athey and Imbens (2015).

References

Angrist, J. D. and Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

Athey, S. and Imbens, G. W. (2015). Summer Institute 2015 Methods Lectures. In *NBER*, Boston.

Caplin, A. and Schotter, A. (2008). *The Foundations of Positive and Normative Economics*. Oxford University Press.

Chang, A. C. and Li, P. (2015). Is economics research replicable? Sixty published papers from thirteen journals say "usually not". *Finance and Economics discussion series*, 83.

Chetty, R. (2013). Yes, Economics Is a Science. *The New York Times*, (http://www.nytimes.com/2013/10/21/opinion/yes-economics-is-a-science.html).

Chetty, R. (2015). Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review*, 105(5):1–33.

Christensen, G. and Miguel, E. (2016). Transparency, Reproducibility, and the Credibility of Economics Research.

Deaton, A. (2009). Instruments of Development.

Deaton, A. and Cartwright, N. (2017). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine*, pages 1–20.

Fama, E. F. and Thaler, R. (2016). Are Markets Efficient? *Chicago Booth Review*, (http://review.chicagobooth.edu/economics/2016/video/are-markets-efficient).

Feyerabend, P. (1975). *Against Method*. New Left Books.

Feynman, R. (1964). The Scientific Method. *Messenger Lectures*, (www.youtube.com/watch?v=EYPapE-3FRw).

Feynman, R. (1974). Cargo Cult Science. *Caltech 1974 Commencement Address*, (http://calteches.library.caltech.edu/51/2/CargoCult.htm).

Friedman, M. (1953). The Methodology of Positive Economics. *Essays in Positive Economics*, II:3–35.

Friedman, M. and Savage, L. J. (1948). The Utility Analysis of Choices Involving Risk. *Journal of Political Economy*, 56(4):279–304.

Gilboa, I., Postlewaite, A., Samuelson, L., and Schmeidler, D. (2014). Economic Models as Analogies. *Economic Journal*, 124(578):513–533.

Glimcher, P. W. and Fehr, E. (2014). *Neuroeconomics: Decision Making and the Brain*. Elsevier.

Goldman, A. I. (1988). *Epistemology and Cognition*. Harvard University Press.

Heckman, J. J. and Urzúa, S. (2010). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156(1):27–37.

Hume, D. (1738). *A Treatise of Human Nature*.

Imbens, G. W. (2010). Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 48(2):399–423.

Keane, M. P. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics*, 156(1):3–20.

Krugman, P. (2013). Maybe Economics Is A Science, But Many Economists Are Not Scientists. *The New York Times*, (http://krugman.blogs.nytimes.com/2013/10/21/maybe-economics-is-a-science-but-many-economists-are-not-scientists/).

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Leamer, E. E. (1983). Let's Take the Con Out of Econometrics. *American Economic Review*, 73(1):31–43.

Leijonhufvud, A. (1973). Life Among the Econs. *Western Economic Journal*, 11(3):327–337.

Leshem, D. (2016). Retrospectives: What Did the Ancient Greeks Mean by Oikonomia? *Journal of Economic Perspectives*, 30(1):225–238.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard University Press.

Manski, C. F. (2007). *Identification for Prediction and Decision*. Harvard University Press.

Marian, D. (2016). The Correspondence Theory of Truth. *The Stanford Encyclopedia of Philosophy*.

Nakamura, E. and Steinsson, J. (2018). Identification in Macroeconomics.

Pfleiderer, P. (2014). Chameleons: The Misuse of Theoretical Models in Finance and Economics.

Popper, K. (1959). *The Logic of Scientific Discovery*. Routledge, 2nd edition.

Romer, P. (2015). Mathiness in the Theory of Economic Growth. *American Economic Review: Papers & Proceedings*, 105(5):89–93.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701.

Rubinstein, A. (1998). *Modeling Bounded Rationality*. MIT Press.

Rubinstein, A. (2012). *Economic Fables*. Open Book Publishers.

Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica*, 5(17):61–71.

Searle, J. (1999). *Mind, Language and Society: Philosophy In The Real World*. Basic Books, 1st edition.

Shiller, R. J. (2013). Is Economics a Science? *Project Syndicate*, (https://www.project-syndicate.org/commentary/robert-j–shilleron-whether-he-is-a-scientist).

Smith, N. (2015a). A paradigm shift in empirical economics? (http://noahpinionblog.blogspot.com/2015/06/a-paradigm-shift-in-empirical-economics.html).

Smith, N. (2015b). Econ critique scorecard. (http://noahpinionblog.blogspot.com/2015/10/econ-critique-scorecard.html).

Smith, N. (2015c). How 'Mathiness' Made Me Jaded About Economics. *Bloomberg View*, (http://www.bloomberg.com/view/articles/2015-05-22/how-mathiness-made-me-jaded-about-economics).

Smith, N. (2015d). Lazy econ critiques. (http://noahpinionblog.blogspot.com/2015/10/lazy-econ-critiques.html).

Thaler, R. H. (2016). Behavioral Economics: Past , Present , Future. *American Economic Review*, 106(7):1577–1600.

Wang, A. Y. (2013). No, Economics Is Not a Science. *The Harvard Crimson*, (http://www.thecrimson.com/article/2013/12/13/economics-science-wang/).

Wigner, E. P. (1960). The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Communications on Pure and Applied Mathematics*, 13(1):1–14.