The conceptual building blocks for making investment scientific

Sergio Focardi, PhD

Professor at University of Genoa, DIME

Founder of Economics and complexity

Causal factor investing has gained recognition as a hot topic in the world of asset management and finance. The book Causal Factor Investing: Can Factor Investing Become Scientific? Authored by Marcos M. Lopez de Prado, published online by Cambridge University Press, October 2023 deserves the credit for directing the attention of the community of investors and researchers to this question. Causal factor models are a fundamental improvement of factor-based asset management. However, in this book, the author has a more ambitious goal: demonstrate that factor-based asset management can become scientific.

In this short commentary I will try to make a contribution by discussing a number of key points about science and the relationship of causation to science that are probably not well known.

Is science unified?

Science is not a unified subject. The notion of reductionism, that is the notion that knowledge can be reduced to a small number of universal laws and principles[1], is no longer widely accepted. In his 1972 paper, More is different, Philip Anderson, winner of the 1977 Nobel Prize in physics, opposed reductionism and claimed that science is stratified in a hierarchical structure. This view is now accepted by many scientists.

There are different types of science. Basic science studies fundamental principles and laws of nature, while different types of science study systems including complex systems, social systems and biological systems. Applied science deals with practical applications of science. Each level in the hierarchy has its own principles. This does not mean that basic laws are violated but it implies that basic laws are probably insufficient to explain complex phenomena. For example, we do not have an explanation of the formation and growth of living objects.

Scientific methodology

Many textbooks and papers claim that science consists in collecting data and then trying to find regularities in the data. This is a fundamentally wrong perception of the scientific methodology. Pure data do not exist, data are always theory laden. Science uses abstract variables that acquire meaning through the entire theory and through observation processes that depend on the theory itself. For example, the enormous success of Newton's Dynamics was due to the simultaneous discovery of the laws of dynamics, including the second law $F = ma$ and the law of gravitation $F = G\frac{m1m2}{r2}$. Mass and force are abstract terms implicitly defined by the theory.

[1] There are different types of reductionism : Ontological, Methodological, Theoretical, Scientific

Science is empirical. Scientific laws are hypotheses that are verified empirically through observations and experiments. The philosopher David Hume pointed out that no finite set of observations can justify generalization to all possible observations. Science is therefore hypothetical.

There is no constructive methodology of scientific discovery. Scientific progress is the product of human genius. Thus far we have not discovered any mechanism for producing new theories. We try to imitate innovation with random searches.

Empirical verification

Scientific laws are hypothetical because they can be verified only on a finite number of observations. In principle scientific laws should be empirically verified on every known fact. According to the Duhem-Quine Thesis theories can be verified only globally. That is, all statements of a theory must be simultaneously verified.

In 1962 Thomas Kuhn argued that scientific progress moves in discrete jumps. As new observations do not fit with existing theories science goes through a paradigm shift which is a fundamental change of the laws as well as of the descriptive framework. In general, the new theories do not explain old observations but create an entire system of new observations. Using the terminology of Kuhn, new theories are generally *incommensurable* with old theories because they use a different descriptive framework.

Demarcation and falsifiability

The demarcation problem consists in separating science from non-science. As science is hypothetical, the philosopher Popper proposed that falsifiability is the criterion for demarcation. In principle a theory can be falsified on a single observation and therefore falsifiability can be proved. Actually, reality is a bit more complicated. Suppose that a new theory is proposed with a falsification procedure. The fact that the theory passes the falsification test does not mean that it must be accepted. The new theory should also pass all known empirical tests. Falsifiability is an important criterion, but it serves only as a conceptual tool for demarcation. In a way or another, theories must prove that they work. This step is never conclusive, but it is necessary.

The notion of demarcation does not imply causality. We can falsify a theory formulated in terms of differential equations. In addition, falsification is subject to the same Duhem-Quine principle: only a global theory can be falsified. The concept of falsification of modern theories is complex. We accept a theory if it explains observations, we reject a theory if it does not explain observations. But both acceptance and rejection refer to global theories.

Science and causation

Causation is a term with many definitions. In addition, it is used in many contexts. Therefore, there is always the risk of misunderstanding. For instance, the book previously cited, Causal Factor Investing, states that a variable X causes a variable Y if Y is a function of X: $Y = f(X)$. However, the book makes it clear that causation is an asymmetric relationship, so that X causes Y but Y does not causes X. In fact in many sentences, it writes $Y \coloneqq f(X)$ which denotes an asymmetric relationship which is not invertible. In fact, the SCM models are written as sets of asymmetric relationships: $F = \{Xi = fi(PAi, Ui)\}$, where the Ui are mutually independent random variables, PAi is the set of variables that cause Xi, and the set of equations F induces a causal graph.

A functional link, per se, is not the signature of a causal relationship. Consider, for instance the classical Newtonian gravitation law $F = G\frac{m1m2}{r2}$. The gravitational force, an abstract variable, depends on the square of the distance between two masses. However, we do not say that gravitation causes gravitational force. The laws of physics describe the evolution of variables. Laws of physics are observational laws that describe the actual real dynamics.

Scientific explanation

At the level of basic laws, scientific explanation is logical not causal. We explain the behavior of physical systems as logical deductions from basic laws plus initial and boundary conditions. This is the Deductive-Nomological principle proposed by Carl Hempel. Therefore, we can say that basic physics is not causal. However, science might study systems that are intrinsically causal. Biological systems, for instance, might exhibit true causal features. So do economies and social systems.

Correlation is not causation

Correlation is not causation is the mantra of causal modelling. But what does it mean in practice? The point is the following. Correlation is observational in the sense that correlation represent regularities that we observe. For example, consider two time series W and V. The two time-series are correlated if positive variations of X are typically associated with positive variations of Y and viceversa for negative variations.

Correlations is an observational phenomenon. We observe that in our samples changes of one variable are typically associated with changes of the same sign of the other variable. But this does not mean that if we force a change of X it is likely that Y will move in the same direction. In this sense correlation is not causation because correlation does not allow to intervene on X to produce changes of Y.

Causation, on the other hand, is effectively the ability to make Y change by intervening on X. If Y is a function of X, Y=f(X), we can say that X causes Y. But causation is an asymmetric relationship: if X causes Y, Y does not cause X. This is why we can say that physics is not a causal science.

Causal mechanisms

Most presentations of causal models claim that causal relationships are produced by causal mechanisms. Let's remark that causal mechanisms are abstract. In most practical cases, causal mechanisms are based on the basic laws of physics. The braking system of a cars is a causal mechanism that ultimately hinges on basic laws. A causal relationship can be expanded in a structure of sub causal relationships. However, every step is purely abstract, there is no ontological commitment.

Can Economics and Finance Theory become scientific?

The answer is probably yes. Most likely Economics and Finance Theory will not become axiomatic theories. Currently mainstream models such as Dynamic Stochastic General Equilibrium Models have the form of axiomatic theories but are very far from empirical theories. They represent idealized economic structures. However, it is likely that economics will become a theory with empirical content. It will be necessary to change the descriptive framework not only dynamic laws.

Causal Factor Investing

The thesis of the book Causal Factor Investing is that a fraction of econometrics is causal and, in particular, a fraction of factor modelling is causal modelling. This is because many econometric models have the mathematical structure of causal models formed by structural equations that are not reversible.

The point is that many econometric models are estimated with least squares methods. Consider, for example, a simple regression. Estimation of the regression with the usual least squares method makes a linear regression an instance of causal models. In fact, estimating the regression parameters with least squares makes the regression equation non-invertible.

The book gives convincing reasons that causal relationships are more robust than relationships based on associations and that estimating causal models as if they were association models is a source of poor out-of-sample performance. Investors should build their investment strategies on causal relationships. This is the main practical message.

The idea that science is causal, and that falsifiability is the basic demarcation criterion should be considered in the context of the universe that we consider. In general, science is not causal. At the beginning of the 20th century philosophers including Ernst Mach and Bertrand Russell strongly rejected causality. There is substantial agreement that basic science is not causal. Basic laws such as Newton's dynamics or Maxwell laws of electromagnetism are not causal, at least not in the sense that causation means manipulability. These laws are purely observational.

However, when we study complex systems such as economies or social systems we might accept that laws are at least partially causal.