Causal Depth: Aspects of a Scientific Realist Approach to Causal Explanation contra Humean Empiricism

Khan, Haider

GSIS

2008

Online at https://mpra.ub.uni-muenchen.de/8293/
MPRA Paper No. 8293, posted 17 Apr 2008 13:23 UTC
Causal Depth: Aspects of a Scientific Realist Approach to Causal Explanation
contra Humean Empiricism

Haider A. Khan
Graduate School of International Studies
University of Denver
Denver, Colorado, USA
hkhan@du.edu
Dec. 2007
Revised 2008
ABSTRACT

The purpose of this note is to clarify how the idea of "causal depth" can play a role in finding the more "approximately true" explanation through causal comparisons. It is not an exhaustive treatment but rather focuses on a few aspects that may be the most critical in evaluating the explanatory strengths of a theory in the social sciences. It presents a general argument which is anti-Humean on the critical side and scientific realist on the positive side. It also elucidates how explanations in political economy and other social sciences can be judged by the scientific realist criterion of causal depth by an extensive example from research in the political economy of development. In this case, an "intentional" and methodologically individualist neoclassical explanation is contrasted with a "structural" dual-dual approach as rival theories purporting to explain the same set of phenomena.

Keywords: Social Explanation, Causal Depth, Scientific Realism, Political Economy, Neoclassical Economics, Structuralism, Social Science Theories, Economic Models
1. Introduction: Causal Depth contra Humean Empiricism

One of David Hume's great insights was to recognize that while events were observable; their putative underlying "causes" were not observable. What were observable such as spatial contiguity, temporal succession etc. turned out in his analysis not to be causes at all but the prejudices of our thinking? The incisiveness of Hume's empiricist analysis which found that there was no causality beyond the empirical regularities observed led Kant to declare that it was Hume who awakened him from his "dogmatic slumber".

Kant's awakening produced the classic "Critique of Pure Reason" where Kant attempted to rescue causality by arguing that by his transcendental deductions causality along with universality, necessity etc. needed to be conceptualized as "categories of our mind". In the twentieth century, the logical positivists rejected the Kantian solution and embraced the Humean position albeit with typical flourishes of symbolic logic that was developing at an accelerated pace in the 1920s. Thus logical positivists such as Hempel or Carnap eschew causal language wherever possible. If cause is mentioned at all, it is to be understood as part of an empirical law or law-like statement. Hempel's deductive nomological model is the most rigorous and elegant example of this practice.

In contrast to the empiricist tradition, the scientific realists have always--- although not always equally clearly or forcefully--- emphasized causal explanations. In the 19th century, Darwin himself was an exponent of this view as a close study of his 1836-1844 notebooks reveal. Huxley was, as usual, much more forthright and pugnacious than Darwin. For example in Huxley (1894:Vol.1;158-9), one reads:

any one who is acquainted with the history of science will admit, that its progress has, in all ages meant, and now, more than ever, means, the extension of the province of what we call matter and causation, and the concomitant gradual banishment from all regions of thought of what we call spirit and spontaneity.

Likewise, George Romanes, another 19th century realist and a friend of both Darwin and Huxley, writes of Darwinism that it

seeks to bring the phenomena of organic nature into line with those of inorganic; and therefore to show that whatever view we may severally take as to the kind of causation which is energizing in the latter we must now extend to the former. . . .the theory of evolution by natural selection . . .endeavours to comprise all the facts of adaptation in organic nature under the same category of explanation as those which occur in inorganic nature – that is to say, under the category of physical, or ascertainable, causation. (Romanes, 1893:402)

What is significant in all these is that Darwin, as Dennet (1995) has most clearly pointed out, more than any other scientist before him articulates and exemplifies a causal approach to scientific explanation by appealing to both non-observable and deep underlying causal processes. It is probably more helpful to study Darwin's notebooks than even the most articulate realist philosopher of science to see how a search for deep, non-mystical explanation can proceed in a complex, almost uncharted territory.¹ In this, there is a lot for social scientists to learn.

However, even if we grant that the notion of cause is not problematic---it actually is still controversial among the philosophers (See Sosa and Tooley (1993) and Salmon (1998)) although specification of causal mechanisms in terms of non-observable entities whose causal effects are in principle observable will be accepted by most scientists as a practical way to proceed\(^2\)--- the question of how to choose from among a list of alternative causes still remains.

The purpose of this note is to clarify how the idea of "causal depth" can play a role in this endeavor. I will not try to be exhaustive but will rather focus on a few aspects that may be the most critical in evaluating the explanatory strengths of a theory. I will present an example from research in development economics where an "intentional" neoclassical explanation is contrasted with a "structural" dual-dual approach to the same set of phenomena.

Although the present approach is, at least in part, intended to draw attention to the revolutionary contributions of Darwin both to science and the philosophy of science, it is compatible with the contemporary developments including Boyd's (and Harman's) abductive defense of scientific realism as inference to the best explanation and Bhaskar's critical realism.\(^3\) However, in contrast with both and consonant with Miller (1987), I focus attention on the need for drawing out the principles of the relevant realist philosophy of science from within the "topic-specific" practices of particular sciences. The purpose of this essay is not to establish this proposition deductively but rather to illustrate its plausibility by pinpointing the philosophical principles embodied in actual pieces of scientific work in a specific problem area.

2. A (Partial) Characterization of Causal Depth and Some Exemplars

The essential idea of causal depth\(^4\) as a selection criterion for explanations is that among the rival theoretical explanations for a given phenomenon, the deepest explanation---deepest compared to its rivals at the time--- is to be chosen provisionally as the "best"("approximately true") explanation. As Miller (1987:88) puts it:

> Suppose a list of causes fits an appropriate standard causal pattern and accurately describes factors sufficient under the circumstances to bring about the effect in question. It may still fail to explain because those causes lack sufficient depth. Roughly speaking, a cause is too shallow to explain why something occurred if it is just one of the ways in which another cause, as intimately connected with the effect, produced the latter. In the slogan version of the causal theory of explanation, such causes are excluded by the requirement that "underlying" causes be described. Actually, "not underlain" would be the more accurate, but absolutely ugly term, since the question is whether one cause is undermined by another.

---

\(^2\) There are some worries in physics about quantum causality and particularly non-locality after Bell's inequality was used by Aspect to show that non-locality did exist at the sub-atomic level. However, given the existence of "decoherence" at a larger scale level of our world, this may not be such a problem. For the social world there still remains the problem of mind-body identity vs. panpsychism and the role of consciousness and meaning that phenomenological and hermeneutic approaches emphasize. However, I believe that at the present state of our knowledge such issues are not resolvable, and we should be open-minded about alternative modes of explanations in the social sciences. Here, clearly the attention is restricted to the set of causal explanations that can be compared meaningfully.

\(^3\) See Boyd(1973,1992),Harman(1965),Bhaskar(1975), Lawson(1997)

More specifically, a cause, X, helping to bring about Y, is too shallow to explain why Y occurred if a cause, Z, of Y undermines X in one of two ways: (a) If X had not occurred, Y would have happened anyway; Z would have produced some causal substitute for X, bringing Y about in some other way. I will label the depth that X lacks, and Z may have here, "depth as necessity." (b) Z is a condition in which Y arose that caused Y, and caused it, in part, by causing X. I will label this aspect or depth using the label "depth as priority."

In both depth as necessity and depth as priority the Darwinian example insists on a trans-empirical and scientific realist theoretical approach. Such an approach pays close attention to both contemporaneous and historical data but at the same time attempts to imagine in a creative manner various causal connections based on non-observable theoretical entities.

In the Humean or logical positivist way of explaining, the dogma is to stick with the "facts" or ---to use a somewhat refined philosophical language--- sense-data. However, as Quine, Putnam and many others have pointed out the Humean rejection of "metaphysics" is itself dogmatic and close-minded since it fails to acknowledge its own ontological commitments.

In some formulations, the Humean "refusal" of ontology can indeed become self-refuting in the same way that relativism is self-refuting. A causal approach as Darwin in particular uses it will sift through a number of rivals to settle upon the most plausible---one could loosely use the language of verifications here as well---alternative in light of the data. Darwin's own example involving bats is particularly elegant.

I have carefully searched the oldest voyages, but have not finished my search; as yet, I have not found a single instance, free from doubt, of a terrestrial mammal (excluding domesticated animals kept by the natives) inhabiting an island situated above 300 miles from a continent or great continental island; and many islands situated at a much less distance are equally barren. . . . Though terrestrial mammals do not occur on oceanic islands, aerial mammals do occur on almost every island. New Zealand possesses two bats found nowhere else in the world: Norfolk Island, the Viti Archipelago, the Bonin Islands, the Caroline and Marianne Archipelagoes, and Mauritius, all possess their peculiar bats. Why, it may be asked, has the supposed creative force produced bats but no other mammals on remote islands? On my view this question can be easily answered: for no terrestrial mammal can be transported across a wide space of sea, but bats can fly across. [More specifically, bat transport occurs to provide a basis for speciation through natural selection, but occurs so infrequently that variants on remote islands are not overwhelmed by migrants from the more competitive mainland.]

It is important to emphasize here that Darwin is engaging in a particularly significant type of causal comparison in this example. As Miller points out:

Here as throughout the book, Darwin is comparing his favored hypothesis of speciation through natural selection not with the mere supposition of its falsehood but with rival hypotheses about the factors at work in the phenomena. The existence of islands with terrain hospitable to terrestrial mammals lacking such endemic species is important because the main rival is the hypothesis, dominant among the best-informed secular-minded scientists of the time, that species are created, without ancestors, by a force that makes them well-adapted to their environments. Also, Darwin makes his argument on the basis of principles he shares with the other side, for example, the shared principle that offspring are like their parents but subject to small variations, not the tendentiously anti-creationist, though plausible principle that a complex organism must be the offspring of another. Finally, Darwin is

---

5 Quoted from The Origin of Species in Miller(1987) p. 164; See also Darwin(1871,1883)
not claiming to have a complete explanation of the phenomena in question, although he certainly thinks that the complete answer would entail the approximate truth of his natural-selection hypothesis. Elsewhere, he makes it clear both that the mechanisms of heredity and variation are mysterious to him and that there is no way of predicting how an observed advantage will affect the actual course of speciation. The issue for him is whether the best available account of the data, however vague or incomplete, entails the superiority of the natural selection hypothesis over its current rivals.6

This type of argumentation is also common in physics when the contest among rival theories demands causally explicit comparisons. An example is Newton's contrast of the causal mechanisms in his celestial mechanics with rival accounts such as Cartesian vortex theories and Tycho Brahe's system.

His discussion of comets in his summary pamphlet, "The System of the World", makes this clear. Indeed, comets are most appropriate for the purpose of causal comparison in this context. All the rival theories in this example share the principles of geometric optics. By using these non-controversial shared principles Newton could derive important features of the orbits of comets.

Once these orbits are derived mathematically, it then can be argued that the Ptolemaic celestial spheres found in Brahe's descriptive geocentric theory can not really exist. For if they did, then surely comets would collide against them. Likewise, the Cartesian vortexes can not be the agents that move the planets and other celestial bodies. Newton observes that comets follow a dynamic trajectory through all parts of the sky which is inconsistent with the dynamics resulting from a vortex.

One could easily multiply such examples from the mature sciences. I hope the above examples demonstrate the need for taking rival theories seriously and for establishing alternative causal mechanisms which can be examined by using techniques of observation and logical procedures which are accepted as fair by all sides. It should also be noted that this is not simply a plea for mainstream theorists to take seriously the causal mechanisms postulated by the non-mainstream theories. The argument is symmetric. In fact, there are grounds for stressing that the challengers to the mainstream theories need to spell out the causal mechanisms of both their own alternative theories and the rival mainstream theories as well as the shared principles among these alternatives. Much dogmatism in social sciences can be avoided if rival theorists were to make explicit the causal mechanisms and the grounds for what would comprise a fair causal comparison among rival theories.

It needs to be emphasized that even after clarifying shared principles, there will generally be substantive areas of disagreement among contending theories. However, in this instance at least, the discussion of substantive disagreements and their possible resolution can proceed without talking at cross-purposes. There are more difficult cases where the framework principles themselves are in dispute. For such cases, it is necessary to develop a detailed theory of confirmation that would rely on topic-specific rules within a field or sub-field of inquiry rather than some global a priori or deductivist general rule (for example, the failed logical positivist attempt to offer such global rules of confirmation for all sciences).

---

While such a theory of confirmation for economics, political economy or the social sciences in general is yet to be fully developed, the approach defended here would call for a consideration of specific debates in substantive areas in order to develop such specific principles of confirmation. In particular, the demands for causal depth in specific theories would have to play a critical role in developing these topic-specific principles.

I now turn from the biological and physical sciences to a topic-specific example in political economy of development. In recent years the neoclassical optimizing agent approach to institutions has gained prominence. The claim by its proponents (see Ray 1997, for a particularly lucid and rigorous presentation of the strongest version of the neoclassical approach) is that these intentional explanations involving the decision making problem of the representative agent are to be preferred to more structural models which do not seem to have "microfoundations". The mathematics of optimization gives it additional credibility. Yet, what are the claims of the more structural "dualistic" models of development inspired by classical political economy?

The most advanced and therefore the most appropriate candidate for causal comparison with the neoclassical model is what can be called a “dual-dual” model (Svejnar and Thorbecke 1980, 1982, Khan 1982a, b, 1985, 1994, 1997, 2004a, b, 2006; Khan and Thorbecke 1988, 1989; Thorbecke, 1992, 1994; Thorbecke and Santiago, 1984; Thorbecke and Morrissom. 1989). This corresponds to the characteristics of a developing economy with not only the traditional and modern sectors but also a kind of dualism within each of these sectors in terms of formal/informal dichotomy. More specifically, the process of development for economies moving from the least developed status to a higher level of development may modify the traditional sector further in the direction of a more market-based modern sector while the formal/informal dichotomy is accentuated within both the sectors. This is one of the most important moves theoretically which is consistent with the stylized facts to be explained in the course of economic development without introducing ad hoc assumptions. In fact, the dual-dual theory reveals that even within the category of the least developed economies the theoretical possibility of the uneven development of the formal and informal sectors both in the urban and the rural areas can exist. In recent years, it has indeed been empirically confirmed as well.

Thus, in this theoretical framework, the coexistence and distribution of modern and informal type of activities in both rural and urban areas are taken as basic structural features of the economy in question. The structural dual-dual approach integrates poverty analysis with rural-urban movements in an economy wide setting by endogenizing both migration and intra-group income distributions and the nominal poverty line. Following this line of work leads to the social scientists' being able to assess policy repercussions on both poverty specific to particular socioeconomic groups and on overall national poverty.

To elaborate a bit further, the starting point is the dual economy models of Lewis (1954) and Fei and Ranis (1964). The roots of Lewis' seminal work are in the

---

7 Within an overall trend towards rapid urbanization there can be migration in both directions. This can have important implications for poverty reduction policies, as Khan (2006) shows for South Asia.
8 See Khan (1997) chs. 2 and 3 for a historical survey and a specific intertemporal dualistic model which is used to analyze the conflict between employment and output.
classical political economy. These pioneering efforts, however, could not or did not take into account the co-presence of dualism within each sector of the two sector models of the dual economy. Erik Thorbecke first raised this issue in 1979 during the course of a National Science Foundation interdisciplinary project on technology and development and Svejnar and Thorbecke (1980, 1982) was the first published work on a prototype of dual-dual technology classification scheme. Khan (1982a, b) and Khan (1985) were applications of this scheme to the energy and textiles sectors in South Korea. Khan (1983) raised the issue of linking technological dualism to poverty theoretically, following an early observation of Pyatt and Thorbecke (1976). Khan and Thorbecke (1988, 1989) were further applications of technological dualism to Indonesia. Khan (1999) explores the connections between rural-urban dualism and migration and poverty in South Africa. Khan (2006) explores both rural-urban and reverse migration in a dual-dual model for South Asia.

In the current formulation, a rural/urban dichotomy is combined with traditional/modern technological dualism, leading to a fourfold classificatory scheme. The four broadly defined sectors in this scheme are:

1. *Subsistence agriculture* with traditional labor-intensive technologies, family farms and food crops for domestic consumption;
2. *Large-scale agriculture* producing mostly export crops using capital-intensive technology.
3. *Urban informal* sector defined in an operational manner;
4. *Modern sector* with industry and services in the urban areas.

Poverty analysis in this dual-dual approach can be integrated with migration and various shocks that are important features of the urbanization process in developing countries. Thus, in addition to explaining the standard "development scenario", the dual-dual model explains additional features of development--- particularly the uneven nature of this process in a number of dimensions--- which the neoclassical approach can not. The fact is that while individual agents may or may not optimize, the causally deeper political economy approach can explain a number of observable phenomena that the neoclassical theory can not explain and this is in addition to explaining the same range of phenomena. Furthermore, in the neoclassical version, the observable market segmentation, surplus labor etching developing economies are derived from the optimizing agent setting only by introducing *ad hoc* assumptions about information, implicit prices, discount rates etc. In formal treatments, the equations of the model as in Khan (2006; 2008) show that the so-called micro-foundations at the household and firm levels can also be incorporated in the formalization of this theory. In addition, the formal/informal and rural/urban structural features are also modeled appropriately.

An important general lesson about causal depth can also be drawn from this example. If scientific realism is even approximately true then pinpointing the ontologically real features in an approximate fashion will likely lead to further discoveries. The history of the mature sciences corroborates this view. It turns out that in social sciences, too, causally deeper theories can lead to deeper insights as well as further discoveries. This

---

is another reason to reject the Humean straitjacket. But more importantly, it is also a
good reason to reject dogmatic, *a prioristic* formulas for doing science even when
they carry the attractive label of scientific realism, as they sometimes do. The moral is
to follow what creative scientists do in any field but to do it with an awareness of the
epistemological and ontological assumptions of rival theories.

3. Conclusions:

In this short note I have tried to defend a partial but useful characterization of causal
depth. I have shown that this view is close to the view implicit in the practice of great
scientists such as Darwin and Newton. In contrast to Hume and Humeans, scientific
realism can advocate such a practice through careful comparisons of relevant theories
in a specific area of research. Such causal comparisons are necessary for judging if
the theory in question really does explain the phenomena it purports to explain more
deeplly than the other contenders. In addition, a "causally deep" theory will generally
lead to further insights that are genuinely novel.

However, it is necessary to sound a note of caution here. Indeed, if one is not careful,
there is a danger of making scientific realism as a philosophy of science---even when
it claims to be critical---as *a prioristic* and dogmatic as logical empiricism. The
antidote to such danger is to cultivate a conscious habit of comparing alternative
theories according to specific criteria for causal depth among other things. Such
comparisons are never complete or final which is really another way of looking at
science as an open-ended social process.
References:


