

# Concretization, Explanation, and Mechanisms

**Abstract.** Traditional accounts of explanation fail to illuminate the explanatory relevance of “models that are descriptively false” in the sense that the regularities they entail fail to obtain. In this paper, I propose an account of explanation, which I call ‘explanation by concretization’, that serves to explicate the explanatory relevance of such models. Starting from a highly abstract and idealized model, causal explanations of the absence of regularities are sought by adding complexity to the model or by concretizing it. Whether this process is successful depends on whether the abstractions and idealizations in the basic model succeed in isolating a mechanism, i.e. in representing how it operates when interfering factors are absent. This account is developed in the context of economics and contrasted to those of Daniel Hausman and Nancy Cartwright. I go on to provide an account of how unrealistic models can be used for providing understanding of the way mechanisms work.

**Keywords:** abstraction, idealization, concretization, explanation, understanding, mechanism, tendency, nomological machine, Nancy Cartwright, Daniel Hausman

Frank Hindriks  
f.a.hindriks@rug.nl

# of words: 7,369

# Concretization, Explanation, and Mechanisms

Derivation plays an important role in many accounts of explanation. Among these is, of course, the Deductive-Nomological account of explanation that was proposed by Carl Hempel (1965). In their accounts of explanatory unification, Michael Friedman (1974) and Philip Kitcher (1981) also require that the explanandum be derived from the explanans. Finally, as we will see below, more recent accounts often retain this idea. On most of these accounts, the first premise of the explanatory derivation should be a general statement, usually of a law. Often such a first premise is itself derived from a model that is highly abstract and idealized. Reliance on highly unrealistic models is not much of a problem for explanations that are meant to pertain to laboratory conditions in which the conditions stipulated in the model are approximated. For many explanations concerning phenomena (or facts) outside of laboratories, however, it is. There is nothing in the actual world to which the alleged explanations that are based on such derivations pertain, which means that they are inadequate.

In this paper, I will present an alternative account of explanation, which I call ‘explanation by concretization’. The main goal that this account is meant to serve is that of showing how highly abstract and idealized models can be of direct relevance to explanations of phenomena (or facts) that obtain outside of the laboratory. The basic idea is that the central implication of such a model is to be compared to that of a slightly more realistic model that is formulated by relaxing an assumption of the original model. The factor that figures in that assumption can then be invoked in order to explain why the model implication of the original model does not obtain in the actual world. Explanation by concretization is meant to be especially relevant to sciences or disciplines in which laboratory work plays a relatively insignificant role. My examples come from economics. I contrast my account to the accounts that Daniel Hausman and Nancy Cartwright have provided of explanation in economics.

The first claim that I will defend below is that explanation by concretization does a better job at capturing the explanatory practices in economics than the accounts of Hausman and Cartwright. Secondly, the account reveals how highly abstract and idealized models, “models that are descriptively false”, can play a central role in successful explanations. And finally, its insights can also be used for developing an account of understanding of the way in which mechanisms work. If the original model succeeds in capturing a mechanism, the process of concretization serves to reveal how it functions in different contexts. Such a sequence of models provides understanding of the mechanism. Thus, I will try to answer a

question that, according to Cartwright (2004, 242), is a pressing one: When does “a false model” provide understanding?<sup>1</sup>

## **1. Concretization Versus Derivation**

### **1.1 Hausman on Derivation**

Both Hausman (1990 and 1992) and Cartwright (1999, 2001) employ a nomological conception of explanation in their accounts of economics. The method of explanation that they share with the other conceptions mentioned above is what I will call ‘explanation by derivation’. The basic idea underlying this method is that the relation between (descriptions of) the explanans and the explanandum is one of derivation: (a description of) the latter is logically entailed by (a description of) the former. Hausman and Cartwright also ascribe an important role to the central implications of the models that scientists employ, which are usually statements of regularities. They take them to be *ceteris paribus* laws that serve the role of explanans. These laws are supposed to be causal laws. I will continue this section by explicating Hausman’s account of explanation in economics in more detail. Subsequently, I will introduce my account of explanation by concretization, and argue that it fits better with the way many economists practice their science. We will also see how it makes highly unrealistic models explanatory relevant. The next section is devoted to Cartwright’s account of explanation in economics. In this account, she makes use of her notion of a nomological machine, a notion that is akin to that of a mechanism. This discussion, in turn, provides the point of departure of the presentation of an account of the role concretization can play in achieving understanding of mechanisms.

According to Hausman (1992) neoclassical economists, especially economists who employ (usually partial) equilibrium theories, follow what he calls ‘the inexact deductive method’. Its first two steps basically consist of the formulation of a potential explanation in a way that is very similar to the Deductive-Nomological method. The first step is the formulation of a law statement, and the second the derivation of a prediction of a particular event from the law statement and a statement concerning initial conditions. The last two steps concern confirmation, which is needed for determining whether such potential explanations

---

<sup>1</sup> I introduced the notion of explanation by concretization in my 2007. The comparison with and criticisms of Hausman and Cartwright as well as the proposal of an account of the understanding of the way mechanisms work are the main new elements in this paper.

are successful. The most distinctive feature of the inexact deductive method is its first step, which Hausman describes as follows: ‘Formulate credible (ceteris paribus) and pragmatically convenient generalizations concerning the operation of relevant causal factors.’ (ibid., 222) The generalizations should be law statements. Rather than universal generalizations, they are restricted in their scope: they are ceteris paribus laws, or tendency statements. Furthermore, the ceteris paribus clauses are not fully specified. Some causal factors are left implicit. And it is not known precisely which factors they are. In other words, the extension of the ceteris paribus predicate is unknown. This is why the method is called ‘inexact’.

Hausman argues that the qualified law statements should be reliable, refinable, and excusable. Reliability is a statistical requirement that Hausman explicates as follows: ‘A generalization such as “ceteris paribus all *F*’s are *G*’s” is reliable only if (perhaps after making allowances for specific interferences) almost all *F*’s are *G*’s.’ (ibid., 140; italics of English words added) Refinability is a matter of it being possible for scientists to make the generalization more reliable or reliable in a larger domain by partly replacing the ceteris paribus clause with specific qualifications that are not ad hoc. A ceteris paribus statement that is lawlike is excusable if the following holds: ‘[S]cientists are able to cite the interfering factors except possibly in a few anomalous cases. It should not seem a miracle that the generalization “works” sometimes and fails others.’ (Ibid., 141) Hausman’s discussion of this final condition reveals that, even though – as we saw earlier – they do not know all, scientists should know the most significant interfering factors. It is not obvious that there are any generalizations in economics that meet these conditions. Perhaps the most plausible candidate for a law statement in economics is the law of demand. Hausman formulates it as follows: ‘[Ceteris paribus people demand more of a good at a lower price.’ (1990, 170) It might not be too farfetched to say that it meets the three conditions just discussed. If so, it could be used for explaining particular instances of the regularity mentioned, according to the account presented above.

Hausman succeeds in capturing several important features of the economics he discusses. Economists often make ceteris paribus claims, and, as we will also see below, they tend to have strong opinions on the major factors that might interfere. In spite of this, I think his account of explanation in economics is flawed. It does not provide a very charitable interpretation of their practice. To see why, we have to consider some of the assumptions economists often make (even more so in the period before 1992, the year in which Hausman’s book appeared, than now). These include perfect information, the absence of transaction costs, and constant returns to scale – assumptions of which we have good reason to believe

they are usually false. These assumptions, however, will usually make up part of the content of ceteris paribus clauses. The reason for this is that the generalizations are formulated on the basis of models that depend on these assumptions, and those generalizations can often not be derived from models in which some of these assumptions are relaxed. In other words, the assumptions are idealizations to which the model implications are often quite sensitive. This means that the ceteris paribus statements economists formulate tend to be of little use for the purpose of explanation. There is little reason to believe that the predictions that can be formulated on the basis of the relevant models will hold. What is more, if the predicted event did occur, this would not be made intelligible by the model employed. On that model it should not have occurred because of the presence of an interfering factor. Now, Hausman is well aware of the fact that a lot of evidence suggests that many of the generalizations economists propose are false (1992, 207 and 209). However, he defends their continued acceptance on the basis of the fact that the ceteris paribus clauses may be the culprit of the failure of the predictions (ibid.). This is, of course, correct. The problem is that it leaves us without an explanation of what is in fact observed. A consequence of Hausman's views, then, is that large parts of economics that economists may be justified in holding on to are of little if any explanatory relevance.<sup>2</sup>

A second criticism of Hausman's account is that it fails to make sense of the widespread economists' practice of relaxing assumptions (which is the topic of Hindriks 2007). Economists often start from highly abstract and idealized, or very unrealistic models. But they usually proceed to formulate models that are less abstract or idealized, or more realistic. This might, of course, be in order to be able to derive model implications that are observed rather than not, and that, as a consequence, can be used for the purposes of explanation by derivation. However, as will be illustrated shortly, the implications of these more realistic models often fail to hold as well. This means that the first criticism applies once again – those models will be explanatory irrelevant as well. An additional problem is that the practice mentioned sits uncomfortably with Hausman's claim that economists usually leave the ceteris paribus clauses largely unspecified. By relaxing assumptions and determining how

---

<sup>2</sup> In a criticism of Cartwright's views, Elgin and Sober (2002) defend explanations that differ from subsumption views of explanation such as Hausman's in that the explanantia involve harmless idealizations that only have negligibly distorting effects on the values of the predictions (as they note, strictly speaking, the explanantia do not entail the explananda in such explanations). Although this move may work for some explanations, it will not work for the bulk of the ones under discussion. They involve gross abstractions and idealizations with non-negligible effects.

sensitive the model implications are to those assumptions, they in effect give further content to the clauses.

The core of the problem surfaces once we realize that the practice results in the production of a large number of models and that Hausman is committed to regarding all the implications of the more realistic models as tendencies. After all, there is little reason to think that, if the implication of the original model meets the requirements laid out above, the implications of more realistic models would fail to be lawlike, reliable, refinable, or excusable. This means we are left with a proliferation of tendencies, which is unattractive for two reasons. First, it trivializes the notion of a tendency. (Note that it also sits uncomfortably with the waning interest in laws in the philosophy of science.) Second, it does not fit with economics as practiced. Not even economists themselves believe they have uncovered many statements worth calling a law (and even Hausman uses scare quotes when talking about what one might call the laws of equilibrium theory; 1992, 209). A related problem is that, interpreted in this way, the practice mentioned would entail that economists are often engaged in documenting a set of tendencies pertaining to one and the same phenomenon without this having a clear explanatory point.<sup>3</sup>

## **1.2 Explanation by Concretization**

According to the alternative view that I will now introduce, the process of formulating more realistic models is of central importance to an important class of explanations. Following Nowak (1989) and Cartwright (1989), I will call the relaxation of an abstracting or idealizing assumption ‘concretization’. In order to explain by means of concretization, one needs to consider two models, the one slightly more realistic than the other. More specifically, the former should be formulated by relaxing one of the assumptions of the latter. As we saw above, the implication of the original model often does not hold in practice. This implication can usually not be derived from the more realistic model. However, it may well be that its own model implication is not observed either. At this point, one might despair as to how the models can be used for explaining anything. A solution comes in view once we let go of the idea that model implications should be regarded as explanantia. Instead, the factor that figures in the assumption that is relaxed can be regarded as the explanans. And the explanandum can consist of the fact that the implication of the original model fails to hold. Let me elaborate.

---

<sup>3</sup> In section 2 we will see that doing so might serve to provide understanding of the underlying mechanism.

Explanation by concretization takes the contrastive approach to explanation as its point of departure (Lipton 1990). Hence, the relation between the explanans and the explanandum is not one of derivation. Instead, an explanation consists of an answer to a question, more specifically an answer to a contrastive why-question. Such a question has the following form (using '*f*' for facts and '*c*' for foils or contrasts): Why *f* rather than *c*? The underlying idea is that we never explain facts simpliciter. The famous example is, of course, that of Willy Sutton the bank robber. When asked by a priest why he robs banks, he answers by pointing out that that is where the money is. This indicates that he has not properly understood the contrast implicit in the priest's question. Whereas Willy took the question to be 'Why rob banks rather than rob something else?' the intended question was 'Why rob banks rather than not rob anything?'. The novelty of explanation by concretization resides in the way the explanans and the explanandum are specified. The fact to be explained is that the model implication of the most unrealistic model of the two used for generating the explanation fails to obtain, as it does in the relatively realistic model and in reality. The contrast is that it does obtain, as is implied by the original model. Using '*r*' for regularities the schema for the explanandum, the contrastive why-question, then, is this: Why does *r* fail to obtain rather than not? The explanans consists of the claim that this is due to the factor that occurs in the assumption that is relaxed. Thus, this method is aimed at explaining the absence of regularities rather than their presence.<sup>4</sup>

As I do more extensively in my 2007, this account can be illustrated using the Modigliani-Miller theorem in financial economics, one of the most well established (in the sense of most widely accepted) theorems in economics at large. According to this theorem, the value of a firm is independent of its financial structure, the debt to equity ratio.<sup>5</sup> Since the introduction of their model in 1958, many models have been formulated in which this implication fails to hold. Modigliani and Miller (1963) proved this themselves by relaxing their assumption that taxes are absent. A differential tax treatment of bonds versus shares that

---

<sup>4</sup> This account of explanation has been inspired by Marchionni's (2006) work on explanation and unrealistic assumptions in economic models. She uses the contrastive approach to explanation in order to argue that different theories that appear to explain the same thing often do not do so.

<sup>5</sup> This is not a statement of a regularity. However, it can be reformulated in terms of the cost of capital: the average cost of capital is uncorrelated with the financial structure of a firm. This is a statement of the absence of a regularity, which can in fact play the same role in explanation by concretization as regularities do. Note that this does not hold for covering-law accounts of explanation, which, therefore, are hard pressed to account for the prominence of the Modigliani-Miller theorem in financial economics.

favors the former over the latter, which is quite common in practice, implies that it is optimal to finance firms with debt only. Kraus and Litzenberger (1973) showed that if, in addition to such a differential tax treatment, bankruptcy costs are present – and they clearly are – there will be an optimal debt to equity ratio for each firm. Jensen and Meckling (1976) in turn have argued that this optimum is influenced by agency costs due to the presence of asymmetries in information (and that the value of a firm is influenced by the extent to which shares are manager-owned).

How can these models be used for explaining anything? Given the account presented above, we should first identify the fact and the contrast in order to be able to formulate the contrastive fact that is to be explained. The fact is that the Modigliani-Miller theorem fails to hold in practice. The contrast is that it does hold. So, the contrastive why-question to be answered is this: Why does the value of a firm depend on its financial structure rather than it being independent from it? This can be explained by means of concretization if there is a more realistic model, one that can be derived from the original one by relaxing one of its assumptions, in which the fact does not obtain. As we saw, this holds for the model that incorporates a differential tax treatment of bonds versus shares. The fact that this factor is present in practice provides a (potential) answer to the question just formulated. The results of Kraus and Litzenberger on the one hand and of Jensen and Meckling on the other reveal that the same holds for bankruptcy costs and information asymmetries. This implies that our explanatory question has no less than three (at least potentially) correct answers.<sup>6</sup>

It is no accident that the first two assumptions to be relaxed were those of the absence of taxes and bankruptcy costs. Modigliani and Miller already singled these out as being particularly important. Although initially they hoped their theorem would represent (the absence of) an empirical regularity, it is now conventional wisdom that it did not and that this is what needs to be explained. Paul Milgrom and John Roberts, for instance, write: ‘Something else, besides the simple workings of classical markets, must account for the effect that financial structure seems to have on what investors are willing to pay. ... The Modigliani-Miller (MM) theorem itself directly suggests several possibilities.’ (1992, 458) In a similar vein, Jean Tirole suggests that the theorem acted as ‘a benchmark whose assumptions needed

---

<sup>6</sup> In my 2007 I point out that each of the more realistic models mentioned has in effect itself been used as the point of departure for another explanation with the same format. The model proposed by Kraus and Litzenberger, for instance, explains why all-debt financing is not optimal rather than being optimal. In other words, it explains why the implication of the model that includes taxes does not hold. Thus, this account of explanation also sheds light on the heuristic value of the process of relaxing assumptions.

to be relaxed in order to investigate the determinants of financial structures' (2006, 1; see also MacKenzie 2006, 89). These passages confirm my claim that explanation by concretization provides a more accurate interpretation of the explanatory practices of partial equilibrium theorists than explanation by derivation (see section 3 of my 2007 for further support). Many economists are aware of the fact that the central implications of their models fail to hold in practice. Nevertheless, they continue to accept them not only because they might be true of situations only realized in possible worlds other than the actual one, but also because they can be used for explaining features of the actual world. All this reveals that highly unrealistic models can be explanatory relevant even if their implications do not obtain.

## **2. Derivation, Concretization, and Understanding**

### **2.1 Cartwright on Explanation**

The main criticism just formulated against Hausman also applies to Cartwright's account of explanation in economics. She defends the idea that economists rely on the method of explanation by derivation, which has the unattractive consequence that, even by their own lights, their models are often of little explanatory relevance. As we saw, this does not fit with the self-understanding of economists. What is more, an alternative account of their explanatory practices is available, which is supported by the way in which many prominent economists talk about their models (again, see section 3 of my 2007 for further support). Furthermore, this method of explanation by concretization enables us to see how highly unrealistic models can be explanatory relevant outside of laboratories. Thus, it can be used not only for making sense of certain explanatory practices, but also for appreciating that such practices may well be valid. The main reason why it is worthwhile to discuss Cartwright's views separately from those of Hausman nevertheless is that she relies on some conception of mechanism in her account of science generally and in that of economics in particular. My criticism of her views provides for a useful point of departure for investigating how concretization can be used for shedding light on the role mechanisms play in the modeling practices of scientists. Sequences of more and more realistic models can provide understanding of the way in which mechanisms work.

As we saw in section 1.1, Cartwright insists that the law statements that scientists employ be causal, and that we should regard ceteris paribus laws as claims about causal powers or capacities (1989, 1999, 2002). More specifically, she maintains that such laws only

hold ‘relative to the successful repeated operation’ of a mechanism, a notion that she partly explicates in terms of laws. For this and other reasons, Cartwright uses the term ‘nomological machine’, which she defines as follows: A nomological machine ‘is a fixed (enough) arrangement of components, or factors, with stable (enough) capacities that in the right sort of stable (enough) environment will, with repeated operation, give rise to the kind of regular behaviour that we represent in our scientific laws.’ (1999, 50) A nomological machine, then, is an entity that has parts and that is located in an environment. Furthermore, it has several capacities that warrant a characterization of the way it operates in terms of laws, as long as certain stability conditions are met. In relation to economics, Cartwright uses the term ‘socio-economic machine’ for what is basically the same idea. I will use the term ‘nomological machine’ instead, because the role that laws play in Cartwright’s conception of a mechanism is of crucial importance to my argument.<sup>7</sup>

Cartwright rejects covering-law accounts of explanation, which form the bulk of the accounts that are instances of the method of explanation by derivation. According to these accounts, explanata are laws or law statements. Cartwright, however, gives priority to natures or capacities over laws. Her alternative to covering-law accounts is explanation in terms of natures, capacities, or nomological machines (ibid., 58 and 138). So, rather than laws or law statements, Cartwright regards capacities or nomological machines or representations of them as explanata. She says little about how exactly this is supposed to work, but we can reconstruct her view by investigating what she says about models and nomological machines. She regards models as blueprints of nomological machines by which she means that models reveal which components such machines have, how they are arranged, which capacities they have, and what behavior, or which regularity results from their joint operation. Cartwright maintains that ‘[w]here there is a nomological machine, there is lawlike behaviour’ (ibid., 57), and that ‘it takes a nomological machine to get a regularity’ (ibid., 73; see also 59). For the case of economics, she formulates this in terms of a slogan: ‘Socio-economic laws are created by socio-economic machines.’ (Ibid., 149) This suggests that, according to Cartwright, we explain regularities rather than singular events, and she does indeed talk of the explanation of regularities (ibid., 58). So, regularities are the explananda rather than the explanantia, as

---

<sup>7</sup> I use the term ‘mechanism’ in a different way than Cartwright. She uses the term for parts of nomological machines, or for capacities (Cartwright 1999, 142-44). On my view, a mechanism is a stable configuration of causal powers. When the term is used in this way, nomological machines are a kind of mechanism.

Hausman has it. And we explain them in terms of models of mechanisms or nomological machines that entail those regularities.<sup>8</sup>

The basic question I want to pose in response, at least insofar economics is concerned, is: Which regularities? The models of economists usually do not imply regularities that are observed in practice. And economists rarely construct laboratory situations in which they might be observed (although they do so more and more). So, economic models rarely capture the actual behavior of economic entities (Cartwright's own criticisms of nomological accounts in fact rely on this claim). Instead, they usually pertain to regularities that could be observed if circumstances were different. This in turn implies that they can at best serve a role in providing potential explanations. Explanation is, after all, factive; in order for an explanation to be successful it must be true. Relative to Cartwright's account of explanation this means that the relevant regularities must actually obtain. As they do not, the kind of explanations Cartwright attributes to economists usually fail. At this point, the argument presented in the previous section can be used again in support of the claim that the account of explanation by concretization is to be preferred as an account of explanation in (significant parts of) economics.

## **2.2 Mechanisms and Understanding**

There is, however, something to be learned from Cartwright's discussion of these issues. The striking thing is that she is, of course, well aware of the fact that few implications of economic models hold in practice. In her discussion of an actual economic model, a game-theoretic model about debt contracts proposed by Oliver Hart and John Moore, she points out that the regularities are derived from the model in the sense of strictly being deduced from it and goes on to note:

The cost is that the rules of the games that allow these strict deductions may seem to be very unrealistic as representations of real life situations in which the derived regularities occur. ... The kind of precise conclusions that are so highly valued in contemporary economics can be rigorously derived only when very special

---

<sup>8</sup> Cartwright (1983) used to subscribe to what she called 'the simulacrum account of explanation'. By and large, the reconstruction I gave of her views on explanation on the basis of her more recent work fits with this account. Perhaps, however, her view on explanation has changed due to a shift in her view on models. She used the term 'simulacrum' because she regarded models as works of fiction (ibid., 153). Now she takes them to be blueprints of nomological machines.

assumptions are made. But the very special assumptions do not fit very much of the economy around us. (Ibid., 148-49; emphasis added)

She does not appear to be worried, however, by the fact that economic models often pertain to regularities that occur only in conditions that are very different from those in the actual world. She writes, for instance, that a model in economics is best represented ‘as a design for a socio-economic machine which, if implemented, should give rise to the behaviour to be explained’ (ibid., 139; emphasis added). It also fits with her talk of models as blueprints of nomological or socio-economic machines. The fact that they are blueprints suggests that they do not (need to) represent the way things actually are.

The solution of this riddle lies in the fact that Cartwright pays little attention to explanation, and focuses instead on understanding. This is apparent from the two problems with regularities in economics that she discusses. The first adds to the problems of Hausman’s version of the method of explanation by derivation as an account of explanation in economics: ‘The most immediate problem with regularities is that, as John Stuart Mill observed, they are few and far between.’ (Ibid., 141) This implies that actual events can rarely be explained by subsuming them under regularities, because there are hardly any regularities under which they can be subsumed. The second problem pertains to the kind of knowledge economists are after:

The second problem with regularities is that, as in physics, most of the ones there are do not reflect the kind of fundamental knowledge we want, and indeed sometimes have. We want, as Mill and Haavelmo point out, to understand the functioning of certain basic rearrangeable components. ... What we need to know is about the capacities of the distinct parts. (Ibid., 141-42)

I think this is exactly right, although we should add that scientists also seek to understand the way in which particular mechanisms work. The next thing we need to know is how models can be used for providing such understanding. This will also serve to answer Cartwright’s (2004) pressing question: When does “a false model” provide understanding?

Cartwright (1999, 53-54) discusses three respects in which scientific understanding of capacities can differ from everyday understanding. First, scientists may be able to ascribe a feature to an entity that is associated with a capacity independently of its display of the capacity described in the related law. Second, they may have formulated an exact functional form of the law that is characteristic of the capacity. Third, they may know some explicit rules for how it will combine with other capacities described by different laws. This third feature

supports the idea to which Cartwright appears to subscribe that understanding of capacities helps to understand how a nomological machine operates. And it is on this that I want to focus in the remainder of this paper. As I am not committed to Cartwright's notion of a nomological machine, I will for this purpose switch to using the term 'mechanism'. I take a mechanism to be a stable configuration of causal powers.<sup>9</sup>

The (or at least a) way to achieve understanding of a mechanism by means of models is to combine derivation with concretization. The first step is to develop a model that isolates the mechanism. Isolating a mechanism is a matter of depicting the way it works in a situation in which no interfering factors are present. This is achieved by means of abstraction and idealization.<sup>10</sup> As a consequence, the model will be very unrealistic. If it succeeds in isolating the mechanism, its central implication will provide a description of the way it functions in the situation modeled. And knowing how the mechanism operates without interference is the first condition for understanding the way it works. I will call a model that aims at isolating a mechanism in this way 'a basic model'. Building such a basic model requires knowledge of the capacities or causal powers involved in the relevant mechanism, and some such knowledge is surely required for understanding the way a mechanism works. The next thing to do is to develop several concretizations of the basic model. The reason for doing so is that one needs to grasp how the mechanism is affected by some major interfering factors. In practice, mechanisms tend to operate in very complex environments. This suggests that only knowing how they operate in the most unrealistic context of all can never add up to genuine understanding. At the same time, however, knowing exactly how it operates in particular cases seems to be too much to ask. Such knowledge will only be available after extensive investigation of those cases. A theoretical scientist need not have such applied knowledge in order to understand the workings of a mechanism. Knowing how the mechanism is affected

---

<sup>9</sup> This definition is not only consistent with Cartwright's conception of nomological machines, but is also compatible, for instance, with the conceptions of mechanisms defended by Machamer, Darden, and Craver (2000) and Bechtel and Abrahamsen (2005), although their explicit formulations are very different. Recall that Cartwright uses the term 'mechanism' differently (note 7).

<sup>10</sup> Mäki (1992) regards abstraction and idealization as kinds of isolation. In contrast to this, I regard isolation as the main goal of abstraction and idealization. I take abstraction to be a matter of omission, and idealization a matter of exaggeration (cf. Cartwright 1989 and Wimsatt 1987).

by some major interfering factors, then, is the second and last condition for understanding in relation to mechanisms.<sup>11</sup>

These two conditions on understanding are directly related to our discussion on Cartwright and explanation in section 2.1. Recall that the criticism was that on Cartwright's account economists provide potential explanations at best. They are not successful because the postulated regularities do not obtain in practice. The thing to see now is that potential explanations are vital for achieving understanding of mechanisms. And these are explanations as Cartwright conceives of them. They are a matter of constructing models of mechanisms that are taken to be the sources of the regularities that would obtain if the appropriate conditions obtained. I suggest calling the mechanism as it occurs in a context without interfering factors, the mechanism as it operates on its own, 'the bare mechanism'. A basic model, then, is a model of a bare mechanism. On the conception I propose and defend in this paper, understanding a mechanism requires understanding the causal powers of which the bare mechanism is composed and their configuration. In addition to this, it involves knowing how the bare mechanism functions and how it's functioning is affected by some major sources of interference. These conditions come down to grasping some potential explanations pertaining to the mechanism in different contexts.

Why should we accept this as an adequate conception of understanding in relation to mechanisms? First, the account resembles an increasingly popular account of understanding in science proposed by De Regt and Dieks (2005) in important respects. De Regt and Dieks defend the following criterion for understanding (CU):<sup>12</sup>

A phenomenon *P* is understood by scientists (in context *C*) if an adequate theory *T* of *P* exists of which they can recognize qualitatively characteristic consequences without performing exact calculations.

Although this requires some unpacking, which I will provide shortly, there is a clear analogy between CU and the condition that a scientist knows how the mechanism functions in various

---

<sup>11</sup> Cartwright (1999) puts a lot of emphasis on shielding conditions, and appears to believe that the functioning of nomological machines can often be easily disrupted. In cases for which this holds, knowledge of such disrupting rather than merely interfering factors should perhaps be required for genuine understanding.

<sup>12</sup> CU results from combining the Criterion for Understanding Phenomena and the Criterion for the Intelligibility of Theories that De Regt and Dieks (2005, 150 and 151) propose. A theory is adequate, according to them, if it meets the usual logical, methodological and empirical requirements (the word 'adequate' is mine).

contexts. The second argument in favor of this account of understanding is that it provides an answer to the question how false models can be used for providing understanding.

Let me illustrate these two points with the Modigliani-Miller theorem. The mechanism that underlies this theorem is what Modigliani and Miller call ‘the arbitrage mechanism’. If it is not interfered with, the outcome of this mechanism as conceptualized by Modigliani and Miller is that the value of a firm is independent of its financial structure. This is a qualitative consequence of the model of the bare mechanism in the sense that De Regt and Dieks have in mind. Significant interfering factors include taxes, bankruptcy costs, and information asymmetries. Again, no exact calculations are needed for recognizing the relevant model implications, which were expressed in purely qualitative terms in section 1.2.

This discussion prepares the way for formulating a criterion for understanding the way mechanisms work (CUM) that is basically an application of CU:

The way a mechanism  $M$  works is understood by scientists (in context  $C$ ) if there is an adequate theory  $T$  about mechanism  $M$  that consists of models of  $M$ , including a basic model, of which they can recognize qualitatively characteristic consequences without performing exact calculations.

Apart from the fact that CU pertains to phenomena and CUM to mechanisms, the main difference between CU and CUM is that the latter mentions models explicitly, while the former only mentions a theory.

Strictly speaking, models are either true by definition, or they are neither true nor false (at least on the semantic conception of theories, which Cartwright also employs). This raises the question what Cartwright (2004, 242) might mean by a model’s being false when she raises the question when a false model provides understanding. Presumably, she has in mind that the claim, or theoretical hypothesis, that the model is descriptively accurate, i.e. that its central implication obtains in reality, is false. This appears to hold at least for the model of the bare mechanism, and those that include taxes, or bankruptcy costs. Still, familiarity with these models and their characteristic consequences lies at the heart of understanding of the arbitrage mechanism in financial markets. Apart from the qualification about terminology made at the beginning of this paragraph, then, CUM explicates how false models provide understanding of the way mechanisms operate. Such understanding presupposes a process of model construction, concretization, and derivation of model implications.

Before concluding, I should stress that mechanisms are also important for explanation by concretization as discussed in section 1. In my 2007 I argue that highly unrealistic models

can legitimately be regarded as true once we realize that they should be evaluated in combination with their assumptions. Formulated more carefully, what I maintain there is that highly unrealistic models such as that proposed by Modigliani and Miller can be used for constructing counterfactuals that are true. The antecedents of those counterfactuals should consist of the assumptions of the models, and their consequents of the model implications. An example is this: If taxes, bankruptcy costs, information asymmetries, and ... were absent, then the value of a firm would be independent of its financial structure. Even though (or if) the value of a firm is in fact significantly sensitive to that structure, this claim may well be true. This line of reasoning has to confront the obvious objection that it is all too easy to construct such counterfactuals. This is why we should insist on the requirement that the most abstract and idealized model of a sequence of models should pertain to (a fundamental force or) a bare mechanism.<sup>13</sup>

### **3. Conclusion**

I have argued that the method of explanation by concretization provides a better account of the explanatory practices in significant parts of economics than the method of explanation by derivation, versions of which Hausman and Cartwright appeal to in their accounts of economics. It does justice to the fact that it is rarely the case that the regularities encapsulated in economic models are observed in practice. This does not detract from their explanatory value, because a highly unrealistic model can be used in combination with a slightly more realistic model for answering why the original model contrast with what we see in practice. Thus, it allows us to see how “false models” can be used for providing true explanations. I continued to argue that such models also provide the (or at least a) key to understanding the way in which mechanisms operate. Crucial to such understanding is knowledge of how a mechanism operates in the absence of interfering factors and of how it’s functioning is affected by some of those factors. Thus, in addition to an account of explanation by concretization, I have defended a conception of understanding by concretization and derivation.

---

<sup>13</sup> I mention this idea in note 20 of my 2007.

## References

- Bechtel, W. and A. Abrahamsen (2005) 'Explanation: A Mechanist Alternative', Studies in History and Philosophy of Biological and Biomediccal Sciences 36, 421-41.
- Cartwright, N. (1983) How the Laws of Physics Lies, Oxford, Clarendon Press.
- Cartwright, N. (1989) Nature's Capacities and Their Measurement, Oxford, Oxford University Press.
- Cartwright, N. (2001) 'Ceteris Paribus Laws and Socio-Economic Machines', in: The Economic World View, U. Maki (ed.), Cambridge, Cambridge University Press, 275-92.
- Cartwright, N. (2002) 'In Favour of Laws that are Not Ceteris Paribus After All', Erkenntnis 57, 425-37.
- Cartwright, N. (2004) 'From Causation to Explanation and Back', in: The Future of Philosophy, B. Leiter (ed.), Oxford, Oxford University Press, 230-45.
- De Regt, H., and D. Dieks (2005) 'A Contextual Approach to Scientific Understanding', Synthese 144, 137-70.
- Elgin, M. and E. Sober (2002) 'Cartwright on Explanation and Idealization', Erkenntnis 57, 441-50.
- Friedman, M. (1974) 'Explanation and Scientific Understanding', Journal of Philosophy 71, 5-19.
- Hausman, D. (1990) 'Supply and Demand Explanations and their Ceteris Paribus Clauses', Review of Political Economy 2, 168-86.
- Hausman, D. (1992) The Inexact and Separate Science of Economics. Cambridge, Cambridge University Press.
- Hempel, C. (1965) Aspects of Scientific Explanation, New York, Free Press.
- Hindriks, F. (2007, forthcoming) 'False Models as Explanatory Engines', Philosophy of the Social Sciences.
- Jensen, M.C. and W.H. Meckling (1976) 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure', Journal of Financial Economics 3, 305-60.
- Kitcher, P. (1981) 'Explanatory Unification', Philosophy of Science 48, 507-31.
- Kraus, A. and R.H. Litzenberger (1973) 'A State-Preference Model of Optimal Financial Leverage', Journal of Finance 28, 911-22.
- Lipton, P. (1990) 'Contrastive explanations', in: Explanation and its Limits, D. Knowles D. (ed.), Cambridge, Cambridge University Press, 247-266.

- Machamer, P, L. Darden, and C.F. Craver (2000) 'Thinking about Mechanisms', Philosophy of Science 67, 1-25.
- MacKenzie, D. (2006) An Engine, Not A Camera: How Financial Models Shape Markets, Cambridge (MA), MIT Press.
- Mäki, U. (1992) 'On the Method of Isolation in Economics', Poznan Studies in the Philosophy of the Sciences and the Humanities, 317-51.
- Marchionni, C. (2006) 'Contrastive Explanation and Unrealistic Models: The Case of the New Economic Geography', Journal of Economic Methodology 13, 425-46.
- Milgrom, P. and J. Roberts (1992) Economics, Organization and Management, New Jersey, Prentice-Hall.
- Modigliani, F. and M. Miller (1958) 'The Cost of Capital, Corporation Finance and the Theory of Investment', American Economic Review 48, 261-97.
- Modigliani, F. and M. Miller (1963) 'Corporate Income Taxes and the Cost of Capital: A Correction', American Economic Review 53, 433-43.
- Nowak, L. (1989) 'On the (Idealizational) Structure of Economic Theories', Erkenntnis 30, 225-46.
- Tirole, J. (2006) The Theory of Corporate Financing, Princeton, Princeton University Press.
- Wimsatt, W.C. (1987) 'False Models as Means to Truer Theories', in: Neutral Models in Biology, M. Nitecki and A. Hoffman (eds.), London, Oxford University Press, 33-55.