The Covariance between Explanation and Prediction in Economic Models

Hanna Schübel*

Abstract

This essay considers the relation between two aims of economics: prediction and explanation. It presents an argument for the covariance thesis, the intuitive claim that predictive and explanatory performances of economic model covary under certain circumstances. The covariance thesis captures the intuition that prediction and explanation are inherently connected: If the relevant independent variables in the model allow for prediction and explanation and the inclusion of more independent relevant variables increases both performances of the model, the performances will covary. A central question of the essay is what constitutes an explanation in economics and different approaches are presented. Finally it is argued that the intuition behind the covariance thesis is in line with Hempel's symmetry thesis.

Introduction

Predictive accuracy has been claimed to be the key epistemic target of economics: "Its [positive economics] task is to provide a system of generalizations

*Many thanks go to Donal Khosrowi for his supervision of this work and to Susanne Burri for her comments on an earlier version.

that can be used to make correct predictions" (Friedman, 1966, p.4). Providing accurate predictions seems what is demanded of economists, particularly when economic or econometric models are used to inform policy formation. What role explanation plays in economics is less clear. In fact, while there are in fact multiple ways to approach the concept of explanation, very little is said by economists on what would constitute an explanation in economics and whether economic models can explain (Hausman, 2001). Yet, especially in the case of policy formation it often seems more necessary to know why some event will occur, rather than merely that it will occur. The thought that explanation and prediction are inherently connected, has a considerable tradition in philosophy: Back in the 1940s it was a significant contribution to the emerging literature on scientific explanation - because of this, more than any other aspect of explanation, it has attracted the interest of many philosophers in the topic (Salmon, 1989).

If we assume that economic models should provide both explanations and predictions - then what capacities of the model allow it to perform these roles and what is their relationship?

In the following I will present a possible answer to this question. I will argue, that the relationship between explanatory and predictive performances of some economic models can be described by covariance. This approach incorporates answers to a number of different questions: What kinds of explanations can economic models provide? What constitutes an explanation in economic models? And what is demanded of a model to provide accurate predictions?

Concerning the economic models my argument applies to, models from econometrics and empirically-informed macroeconomic models are the main target. My paper does not cover more theory-driven models that get their structural specification from (established) theory where only some of the parameters are empirically estimated (since they are known to vary with the context), nor does it apply to models that are exclusively theory-driven. The covariance thesis is limited, in the sense that it applies only in a very thin sense, capturing different approaches to explanation and prediction. At the same time it manages to capture the intuition that prediction and explanation are inherently connected and strongly related.

In what follows I present my argument for the covariance thesis and then enter the discussion on explanation and prediction in economics to see what implications can be derived from the covariance thesis. Specifically, I will turn to the question what kinds of explanations can be provided by economic models, how explanation and prediction in economics are connected with correlation and finally what kind of support Hempel's approach can offer for the covariance thesis.

The Covariance Thesis

In order to argue for the covariance thesis, I will draw on linear regression models. This of course limits the covariance thesis, and more limitations will become clear as we dive into the implications of the thesis. However, I would first like to present my rather intuitive argument for describing this relationship.

Consider the following simple example of a linear model (following Woodward, 1995, p.10): A dependent variable Y is a function of a set of independent variables x_1, x_2 ... x_k and an error term u. As an example take the wage rate Y in a company to be determined by the marginal productivity of the company which is only dependent on the input of working hours x_w (with x_w being part of the set of x_1, x_2, ... x_k) Assume that we have T observations of the values of Y and T corresponding observations of the values of x_1... x_k, via the regression coefficients B_1, B_2 ... B_k. The linear association between the dependent and independent variables could look as follows:

$$Y_i = B_0 + B_1 x_{1i} + B_2 x_{2i} + ...B_k x_{ki}, i = 1, 2...T$$

The main idea is to add more and more variables (x_k) that are relevant for determining the dependent variable Y (e.g. the wage rates of other companies in the sector). This makes it possible to attribute changes in the dependent variable to changes in the relevant independent variables.

The question now is how explanation and prediction can be derived from such a model. By including relevant variables for the occurrence of Y and their effect on the outcome Y in the model, prediction is made possible: The change in a relevant variable can be used to predict the change of the dependent variable. A model's predictive performance is high if the model provides "correct predictions about the consequences of any change in circumstances. Its performance is to be judged by the precision, scope, and conformity with experience of the predictions it yields" (Friedman, 1966, p.4).

It seems intuitively plausible that the model's predictive performance will increase with the number of variables added into the model. For example, if I wanted to predict the development of a wage rate, the inclusion of variables such as the strength of labour unions and governmental legislation would increase the predictive performance. This would increase even more by including, for example, a variable corresponding to the ability of companies to pay their workers. A highly predictive performance in my conception would not only apply to the accuracy of the prediction in a specific setting (under consideration of a fixed group of variables), but the model's ability to make (fairly accurate) predictions in many different scenarios.

At the point where all variables are added, the model will be fully specified, and the predictive performance of the model at its highest. However, this does not necessarily imply that the increase in predictive performance is monotonically related to the number of variables included.

For the explanatory performance of the model, it is important to note that the above model refers to variables that correspond to observations; it links observations of the values of Y (output) with observations of the input variables. It thereby shows a possible relation between variables and possible changes in the phenomenon or dependent variable. The model is then "explanatory" in the sense that it shows how much variation in the dependent variable can be attributed to the variance in the independent variables, without necessarily making statements about the causal relation between them. The explanatory performance of a model then increases with the number of relevant variables included in the model, as more variance can then be attributed.

However, this also leaves us with the important question as to whether explanatory performance of a model understood in this way can help to answer the overall question - that is, what kind of explanation we are after in economic models. These questions will be addressed later. For now, assume we can define explanation as the "attribution of variance". Capturing numerous independent variables (x_1, x_2 ... x_k) and stating their influence on Y would increase both the model's predictive and explanatory performances. An increase in the performances of the model on either of these metrics would be accompanied by an increase or the staying constant of the other performance: The performances both depend positively on the degree of specification of the underlying mechanism by the economic model and the number of the independent, relevant variables included in the model. This indicates that if a model's capacity to predict is high, the model's capacity to explain would be high as well, and vice versa - and if one of the capacities increased, the other capacity would increase as well (though possibly at a different rate).

If this is the case, the relation between the predictive and the explanatory performance of the model can then be described by the covariance thesis: *The relationship between explanatory capacities and predictive performances of economic models is one of covariance.*

This picture is of course oversimplified. In the case of interaction terms between the variables and non-linear associations, these relations would need to be included as well. For the predictive performance of a model to increase when we add new variables to the model, we would have to assume that by adding more and more variables to the model, the invalidity of the model becomes less and less of a problem. Finally at the point where all relevant variables are included, the model will always have zero error. What we can note in favour of the intuitive argument just laid out is that the covariance between prediction and explanation seems to hold at the point where we have included all relevant variables in our model: Under the assumption of a stable mechanism and lack of interventions, we would have a perfect prediction of future events as well as a perfect explanation of past events.

In the following discussion, I want to shed light on some of the implications and possible shortcomings of this intuitive approach. I will discuss the assumptions that the explanatory performance and predictive performance of a model increase by including independent variables. As a main point, I will discuss the different implications that definitions of explanation and prediction have on the predictive and explanatory performances outlined.

Discussion

A. What kinds of explanations can economic models provide?

If we want to quantify the explanatory performance of a model, the question as to what kind of explanations economic models can provide needs to be addressed. A helpful distinction to get a grip on this question is the one between "type" and "token" explanations: "Token" explanations explain specific events, "type" explanations aim at explaining a specific kind of event (Salmon, 1992). This is the "important distinction [...] between explaining singular events (such as the UK's ejection from the ERM [European Exchange Rate Mechanism]) and explaining generalisations, types of events, or states of affairs (such as a stock market crash or the downward rigidity of wages)" (Runde, 1998, p.156). In many economic models what seems to be demanded of an economist is an explanation of a specific event: Macroeconomic models make generalized assertions about the markets and consider types of macroeconomic variables. They aim at explaining types of behaviours of or between economies, just as microeconomics sets out to describe general behaviour of economic agents. But can an economic model provide "type" explanations?

Aiming for explanations of general "types" of events has an effect on the kind of explanation that can be provided since "generalisations do not have unique causal histories" (Runde, 1998, p.156) Explanations for generalizations then consist of descriptions of causal factors "that are common to (many or most of) causal histories of the singular events which fall under those generalisations." (Runde, 1998, p.157) This idea of explaining generalizations by describing causal factors that are common to singular events is interesting, since there may not be just one single explanation of a phenomenon and its general occurrence.

Nevertheless, this leads one to ask which variables should be included in our model to arrive at an explanation. Even previous knowledge about the causal structure raises the substantial problem of deciding which variables play a role in the process of bringing about a phenomenon, and the inclusion of different sets of variables or different assumptions about their relations will change the eventual predictions for the dependent variable Y (Woodward, 1995). The consideration of only some relevant, independent variables, which are by themselves sufficient to bring about a certain phenomenon or event (Persson, 2009), would lead us to "how possibly" explanations. Citing potential causes, which are sufficient to bring about phenomenon X, they leave open the possibility that the phenomenon in a specific event X was caused by some other arrangement of sufficient conditions. How could we quantify the explanatory performance of such a model? Would it be high or low, depending on what variables we identify as having influenced the phenomena, which may - or not may - be the ones referred to by the "how possibly" explanation? Imagine two models, set out to explaining the same phenomenon, yet built on different sets of relevant variables. Both have a high explanatory performance since they refer to variables describing causes that are sufficient in bringing about the event.

Is this then also the case for the predictive performance? Which variables need to be included in order to have a high predictive performance? Consider the example of the financial crisis 2008. Economists cite causes that are possible explanations for the emergence of the crisis and if they know that the causes are sufficient to bring about the event and if they also know about the likely presence of these causes, they might be able to make a prediction of the new occurrence of the event. But this prediction is limited, since occurrence of a new financial crisis may be determined by other factors, which are not captured by the "how-possibly" explanation. A model could also be built on a different set of relevant, independent variables that are included in a different how-possibly explanation. Thus if the future event was caused exactly by the factors named in the "how possibly" explanation, the predictive performance taking these factors into account would be high. But this would imply an exact knowledge of the causal chains leading to future events. Findings such as "how-possibly" explanations can then provide useful knowledge, but do not necessarily lead to a highly predictive performance of the model.

The inclusion of only some relevant variables seems to pose problems for identifying the explanatory performance of models just as for the predictive performance. To have a model with a high explanatory and predictive performance, would we have to include all relevant variables in the model? It seems that in such a scenario, we could describe the relation between the two performances by covariance.

But in events that are in their occurrence influenced by a multitude of factors, such as macroeconomic events, it seems too demanding to include all variables that have had some impact on the event. Conceptually, the notion of "how-actually" explanations may offer us a way around this problem: "How-actually" explanations aim at citing the actual causes for the occurrence of a phenomenon or event X, which means they do not only provide the information that Y is sufficient for X, but also that Y was indeed the case and that Y was suitably involved in bringing about X. If we can identify these variables, it seems that the explanatory performance of the model would both be highest, without identifying all relevant variables. By referring to "how actually" explanations for the specific events would mean to have the highest explanatory performance of a model, even for the generalization of the phenomenon.

The question how we can quantify the explanatory performance of an economic model prevails. I claimed that the simple linear regression model I presented is "explanatory" in the sense that it shows how much variation in the dependent variable can be attributed to the variance in the independent variables. At this point, having a "how-possible" explanation that has these effects or a "how-actually" explanation citing the actual causes for the phenomenon on the way, once more variables were included, does not make a difference for the covariance thesis.

When we ask what kind of explanation economic models can provide, the notion of explanation as attributing variance in the dependent variable to variance in the independent variables seems a not very far reaching account of explanation. Economists may be interested in explanations that "go further" like, for example, why the change in the minimum wage effected a particular change in the unemployment rate rather than some other change. The attribution of variance is a basic concept that can be advanced and serves as a basic argument for the intuitive argument for the covariance thesis.

B. The challenge of correlation

The definition of explanatory performance as that of "attributing variance" is especially useful in the light the following objection to the idea, that there is a strong relation between the predictive and explanatory performances of some economic models. The claim that in order to give an explanation some reference to a causal relation is necessary - while it is not necessary for prediction: "prediction requires only a correlation, the explanation more" (Scriven, 1959, p. 480). For example, consider Okun's Law, an econometric finding that (originally) states that the unemployment rate declines when growth is above the trend rate of 2.25 percent (Dornbusch and Fisher, 1994, p.17). This allows prediction but for no explanation at all.

The claim that explanation requires more than correlation, points at a far reaching discussion of causality in economics and econometrics. Some argue, that correlations can be interpreted as causal relationships in special cases, for example if a relationship between variables remains "stable or invariant under various possible changes or interventions" (Woodward, 1995, p.12) and the correlations shows stability or invariance correlation under some class of changes. Under this approach, explanation is possible since "within the scope of invariance of those claims we can nonetheless legitimately interpret them as causal relationships [...] and we can make them provide explanations" (Woodward, 1995, p. 26). Returning to the objection to the covariance thesis that there are models that provide prediction but no explanation, we can state that in the case where those models refer to correlation, this does not hold. One reason is that we can derive explanation from a specific kind of correlation, namely a very stable one. It seems also worth noting that a prediction referring to an unstable correlation will most likely lack accuracy.

C. Hempel's symmetry thesis

At this point the question needs to be asked whether, under my conception of "attributing variance" for explanation, there is still a difference between explanation and prediction. An important conceptual difference, in line with my idea, is emphasised by Hempel in his symmetry thesis. Hempel argues that depending on "the time of the derivation, the derived data are, or are not yet, known to have occurred, the derivation is referred to as explanation or as prediction" (Hempel & Oppenheim, 1948, p.164). This is referred to as the symmetry thesis: explanation and prediction have the same logical structure and only differ in the temporal aspect. They are "symmetrical in terms of the inductive and deductive inferability" (Caldwell, 1994, p.55).

Hempel's covering-law model (1962), in which the symmetry thesis is set, concerns two kinds of explanations, which are in fact arguments, deductive-nomological and inductive-probabilistic explanations. Both deduce or infer their result from either general laws or probabilistic-statistical laws (Hempel, 1962). In economics "the majority view is that deductivism continues to be the best approach on offer" (Runde, 1998, p. 151). But both kinds of explanation have been found to be "too restrictive in their characterization of what is to count as a legitimate explanation, thereby excluding many types of explanations which are considered legitimate by scientists" (Caldwell, 1994, p.54). One of these restrictions of deductive-nomological explanations is that they refer to deterministic laws, which do not seem to exist in the domain of economics. The explanatory status of economics (together with biology and psychology) is questioned by the deductive-nomological model: "These sciences are full of generalizations that appear to play an explanatory role and yet fail to satisfy many of the standard criteria for lawfulness" (SEP, "Philosophy of Economics"). The model used above would also be excluded: Adding independent relevant variables in order to increase the explanatory performance of a model, without referring to a general law, would not help us derive an explanation.

The basic claims of the symmetry thesis and the covariance thesis are compatible with each other: If we had deterministic laws in economics or full knowledge of the causal factors involved in the production of a phenomenon or full knowledge of all independent relevant variables and their covariance with the dependent variable and full knowledge of the circumstances - then the explanatory and predictive performance of the model would be both at their highest degrees. Hempel's thesis can provide basic support for the covariance thesis and the covariance thesis covers several aspects which the symmetry thesis is not aiming to describe, in that the symmetry thesis refers to a specific kind of explanation.

Conclusion

Vast discussion of what constitutes an explanation - in science and in economics - focuses attention on the various attempts to characterise this aim of economics. The arguments for the covariance thesis I have presented leaves open the question as to what would be a "high explanatory performance" of an economic model. It seems likely that a model would have to take into account relevant, independent variables which are actually involved in specific events in order to arrive at an explanation about generalizations of events. The covariance thesis captures the idea that prediction and explanation are inherently connected. The intuitive argument I have presented rests on the assumption that the predictive and explanatory capacities of the model increase with the inclusion of relevant variables. I have investigated in several different aspects of these assumptions, considering the difficulties of quantifying either of the two performances of a model. It can be noted that in the case of the inclusion of all the relevant independent variables, covariance thesis holds. This argument is in line with Hempel's description of the relation between scientific explanation and prediction.

In this framework it is important to mention that the questions concerning whether economic models need to give a full description in order to fully analyse the causal mechanism behind a phenomenon will have to be answered independently of the question whether economics can in fact provide explanations. Furthermore it seems important to note that even though a model may not provide a "how-actually" or "full" explanation or a perfect prediction it may never the less provide some genuinely useful knowledge.

References

[1] Caldwell, B.J. *Beyond Positivism - Economic Methodology in the Twentieth Century*, revised edition, London, Routledge, 1994.

[2] Dornbusch, R., Fisher, S. *Macroeconomics*, 6th edition, New York, McGraw-Hill, 1994.

[3] Friedman, M. "The Methodology of Positive Economics". *Essays In Positive Economics.* Chicago, University of Chicago Press (1966).

[4] Hausman DM. Explanation and diagnosis in economics. Revue internationale de philosophie. 2001(3):311-26.

[5] Hempel CG, Oppenheim P. Studies in the Logic of Explanation. Philosophy of science. 1948 Apr 1;15(2):135-75.

[6] Hempel, C. G. "Two Models of Scientific Explanation". in "Explanation in Science and History". *Frontiers of Science and Philosophy*, ed. R.G. Colodny, pp. 9-99, Pittsburgh, Pittsburgh University Press, 1962.

[7] Kitcher P, Salmon WC. Minnesota studies in the philosophy of science.

[8] Persson J. Three conceptions of explaining how possibly - and one reductive account. In EPSA Philosophy of Science: Amsterdam 2009 2012 (pp. 275-286). Springer Netherlands.

[9] Runde J. Assessing causal economic explanations. Oxford Economic Papers. 1998 Apr 1;50(2):151-72.

[10] Salmon, W. C. "Scientific Explanation." *Introduction to the Philosophy of Science*, by J. Earman et al. M. Salmon, pp. 7-41, Englewood Cliffs, Prentice-Hall, Inc, 1992.

[11] Scriven M. Explanation and prediction in evolutionary theory. Science. 1959 Aug.

[12] Standford Encyclopedia of Philosophy: "Philosophy of Economics" (revised 2012, derived from http://plato.stanford.edu/entries/economics/)

[13] Woodward J. Causation and explanation in econometrics. In On the Reliability of Economic Models 1995 (pp. 9-61). Springer Netherlands.

Hanna Schübel has just started her studies in the MSc program Politics, Economics and Philosophy at the University of Hamburg (2016). She holds a BA in Philosophy & Economics from the University of Bayreuth. Her main areas of interest are political & legal theory and philosophy of science. You can contact her at [hanna.schuebel@gmail.com]