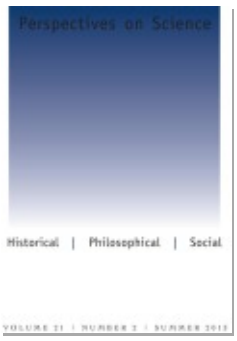
Relations Between Theory and Model in Psychology and Economics

Till Grüne-Yanoff

Perspectives on Science, Volume 21, Number 2, Summer 2013, pp. 196-201 (Article)

Published by The MIT Press



Relations Between Theory and Model in Psychology and Economics

Till Grüne-Yanoff
Royal Institute of Technology, Stockholm

This commentary discusses Nurmi's account of the relation of "theories" and "models," and compares it to methodological discussions in economics. I differentiate between two styles of modeling, "theory first" v. "data first," and argue that Nurmi's recent work falls in the "theory first" category. This implies that models at any level can function as agents of change that may unidirectionally affect models at other levels.

1. Introduction

For Jari-Erik Nurmi, the practice of model-making in psychology is a complex process operating on different levels simultaneously. At first sight, his account seems to reflect Suppes' (1962) notion of a hierarchy of models: from low-level data models to high-level theoretical models, where at each level the model represents "structure" at a different degree of abstraction, and the levels are connected through structural isomorphism.[1]

In this commentary, I want to complement and perhaps somewhat redirect Nurmi's analysis of his own modeling efforts—away from the idea of an interconnected hierarchy of isomorphic structures, towards more autonomous roles of the models at different levels, each with its own intermediate functions.

Focusing on just two of these levels—what Nurmi calls "theories" and "models," I compare Nurmi's account to a related methodological discussion in economics, which distinguishes two modeling styles, "theory first" and "data first." Then I argue that his own recent modeling work falls in

1. Suppes' semantic account of models seems often to be considered the standard account by modelers, not only by Nurmi. Read (in this volume), for example, offers similar views.

Perspectives on Science 2013, vol. 21, no. 2

doi:10.1162/POSC_a_00092

the "theory first" category. This opens an interesting gap between his methodological reflections and his own modeling practice.

2. Theoretical and Empirical Modeling

Nurmi distinguishes between theories and models. Theories express paradigms, traditions and cumulative evidence. They are used to direct research. Models, in contrast, express observations and data. They are used to test and falsify theories.

Nurmi's distinction between theory and model closely matches the distinction in economics between theoretical and empirical (or econometric) model. Typical purposes of theoretical models in economics are the illustration of underlying theoretical principles or hypotheses, and checking the results of their interactions. Econometric models, in contrast, specify the concrete functional forms for estimation.

Interestingly, Nurmi describes the relationship between theory and model (in figures 1 and 2) as a hierarchy of abstraction in which lower level models influence higher-level theories *and vice versa.* I show that this is not necessarily so by drawing on some recent discussions of the methodology of econometrics.

3. Two Styles of Empirical Modeling

The relation of theoretical and empirical models in economics has been the focus of considerable research. In particular, two styles of empirical modeling have been distinguished.

Econometric modeling often takes the form of *theory-first.* In this modeling style, the functional form of the econometric equations are derived from relevant theory, then the parameters of this function are estimated from the sample of relevant data one could obtain, and finally, statistical tests on these estimates are performed in order to assess the adequacy of the empirical model. If these tests turn out negative, various options are possible, including changes in the theory.

In this style, econometric procedures function as direct estimations of theoretically articulated structures (e.g., Hoover 2005, p. 30). The theoretical model is translated into an empirical model, the empirical model tested, and conclusions for the theoretical model drawn from this test. No alternative empirical models are considered. Variable specifications, directions of dependencies, and operationalization of variables are determined by the theoretical model, which ultimately is the main focus of the empirical investigation.

In contrast, econometric modeling is sometimes takes the form of *data-first.* In this modeling style, the modeler searches over different ways to process data statistically with the aim of finding a specification that meets

certain criteria. That is, the data is searched for patterns, and these patterns are then represented in an empirical model. Such a representation is never free from theoretical considerations, as the modeler must make decisions about how data is elicited and measured, and which patterns are to be focused on. But when pursuing the data-first modeling style, it is not a previously developed theoretical model that determines model specification or selection.[2]

This is commonly known and often abhorred as "data mining", but might under suitable regulation be a legitimate procedure. In this style, econometric procedures function as "filters that process raw data into statistics" (Hoover 2005, p. 29). Their results are neither valid nor invalid, but useful if they reveal theoretically interpretable facts about the world. As Hoover comments elsewhere:

> The goal of econometrics is . . . to discover facts that are generated by unobservable nomological machines, facts that theoretical models explain by providing rough drawings, if not blueprints (Hoover 2001, p. 53)

Theoretical models, in this style, thus are not used *prior* to and *for* the construction of econometric models, but rather for the explanation of empirical facts *after* they have been identified by empirical models.

The two different styles thus imply very different relations between theoretical and empirical models, and thus very different meanings for the respective "up" and "down" arrows in Nurmi's figure 1. The theory-first style assigns discovery and construction solely to theory, and only justification to econometrics. The data-first style assigns a good part of discovery and construction to the empirical model. Thus there is no one-to-one, top-down relationship between theoretical and empirical model, but rather a set of empirical models that is related to a theoretical model after that set has been sufficiently analyzed.

4. In Psychology

In this section, I briefly survey three research papers by Nurmi and coauthors, paying particular attention to the relation of theoretical and empirical model.

In Salmela-Aro et al. (2007), the authors present a theoretical model of socialization, with three main implications: Age-graded environments

2. A typical example of this style is found in the so-called LSE method. It proceeds by searching for econometric models that are (i) valid parsimonious restrictions of the completely general (theoretical) model, and (ii) that do not contain an even more parsimonious valid model within them. Thus the theoretical model constraints the search, but does not determine the specification.

channel personal goals, goals affect *selection* of environments and life paths, and people *adjust* their goals on the basis of previous life transitions.

The authors perform a longitudinal study, in which they track young adults' goals, their age and the type of life events they experience over a period of ten years. They analyze the resulting data by estimating Latent Growth Curves models for each of the goal types, and by estimating the dependency of the "level" and "slope" of these goals on age, gender, and the experience of various life events. The resulting parameter estimates then are interpreted as "supporting the [theoretical] model for the most part" (Salmela-Aro et al. 2007, p. 708).

The analysis is driven by the underlying theoretical model of socialization in at least two ways. First, the specification of empirical model and its independent variables are driven by the theory. The authors do not mention testing possible correlations between error terms and the regressors (which would indicate an omitted variable bias) nor do they search for alternative specifications and compare test results of these with the original specification.

Second, their paper discusses two processes going in opposite directions: goals affecting life-paths and life events leading to adjustments in goals. Yet the empirical model only offers correlations between these two factors. No attempts are made to determine the directions of these measured dependencies empirically. Instead, the direction of the measured correlations is interpreted through the lens of the theoretical model.

Nurmi and Salmela-Aro (2002) presents a theoretical model in which the construction of goals is based on a comparison of individual motives with developmental tasks, institutional opportunities and age-related social constraints on the one hand, and on experiences of success or failure in dealing with particular transitions on the other hand.

The authors perform an observational study that recorded personal projects, personal moods and questions regarding work status. They analyze the data in two regression analyses. Using the maximum likelihood estimator, they test five hypotheses implied by the theoretical model. Again, specification of the estimated variables is based on the theoretical model. No comparisons to alternative specifications are offered.

Nurmi et al. (2008) present a theoretical model of the relation of a person's feelings of work exhaustion and her appraisal of goal confidence. The model proposes that goal confidence affects the level of exhaustion felt in the next period, while conversely the level of exhaustion affects the goal confidence at the next period.

The authors perform a longitudinal study, in which they construct the two factors of interest, exhaustion and goal appraisal, from a series of observable variables. They analyze the data with a time series model for each

person individually, regressing each observed variable on the respective lagged latent variable. That is, they investigate the dependency of, say, tiredness, on the level of work exhaustion from the previous period.

To test the fit of their model, they construct a null model, in which each observed variable is regressed on its own latent variable, and then compare their own with the null model through the Akaike information criterion (AIC).[3] No other models than this null model are offered for comparison. Thus again, the authors draw on the theoretical model to specify the regression model, and to determine the relation of observed and latent variables.

Thus in all three papers, Nurmi and collaborators use empirical modeling as measurement of previously established theoretical models. Possible further inferences from the empirical data—e.g., about variable specification or the type of curve—remain unused.

5. Conclusion

Nurmi describes the relationship between theory and model with arrows going both ways. By identifying his modeling style as theory-first, and *not* data-first, I was able to specify these relationships more closely. The arrow from theory to model determines variable specification and shape of the curve. The arrow from model to theory provides estimates of the parameters, possibly offering inferences about the validity of the theoretical hypotheses. Further inferences from the model, however, are *not* included in this arrow.

Thus, *pace* Nurmi's figure 1, the influence of theory on model is *not* symmetric to the influence of model on theory. Rather, the influence of theory on the model is stronger than the other way around, as the theory contains and transmits information that the model does not contain or transmit to the theory. Nurmi himself seems to describe this state in his analysis of example 1:

> The theory . . . provides a heuristic tool for formulating models that can be tested empirically, and, as such, no single study or experiment can either prove or falsify it (Nurmi, this volume, 185–186).

Identifying Nurmi's modeling style as theory-first explains how the theory can function as such a heuristic tool. Instead of being part of a hierarchy of abstraction where each level maintains structural isomorphism,

3. The AIC trades off the sum-of-squares of residuals with the number of regressors, thus safeguarding against the problem of overfitting. In the case described, the null model has a drastically higher number of regressors than the proposed model (5 vs. 2). It thus does not surprise that the proposed model does well in this comparison.

Nurmi singles out the theory as the prime agent of development for his investigation, which integrate various research findings concerning adolescents' future orientation" (Nurmi, this volume, 185–186). It is through this process of integration in the theory (i.e., the theoretical model) that Nurmi commences a new research endeavor, with significant effects for how measurement proceeds, how data is collected, how the functional form is specified and how it is estimated. Instead of being an integral part in a tightly-knit hierarchy, the theoretical model in Nurmi's modeling practice assumes the role of a prime mover.

References

Hoover, Kevin D. 2001. *The Methodology of Empirical Macroeconomics.* Oxford: Oxford University Press.

Hoover, Kevin D. 2005. "Quantitative Evaluation of Idealized Models in the New Classical Macroeconomics." Pp. 15–34 in *Idealization XII: Correcting the Model. Idealization and Abstraction in the Sciences.* Edited by Martin R. Jones and Nancy Cartwright. Poznan Studies in the Philosophy of Science and the Humanities. Amsterdam: Rodopi.

Nurmi, Jari-Erik, Katariina Salmela-Aro, Pertti Keskivaara and Petri Näätänen 2008. "Confidence in work-related goals and feelings of exhaustion during a therapeutic intervention for burnout: a time-series approach." *Journal of Occupational and Organizational Psychology* 81:277–97.

Katariina Salmela-Aro and Jari-Erik Nurmi 2002. "Goal construction, reconstruction and depressive symptomatology in life-span context: the transition from school to work." *Journal of Personality* 70:385–20.

Katariina Salmela-Aro, Kaisa Aunola and Jari-Erik Nurmi 2007. "Personal goals during emerging adulthood: a 10-year follow-up." *Journal of Adolescent Research* 22(6):690–715.

Suppes, Patrick 1962. "Models of Data." Pp. 252–61 in *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress.* Edited by E. Nagel, P. Suppes and A. Tarski. Stanford: Stanford University Press.