Great Economists of the Past

Honors

- Nobel prize (1989)

What he is (most) famous for

- **The Probability Approach in Econometrics**
    - *The statistical implications of a system of simultaneous equations* (Econometrica '43)
    - *The probability approach in econometrics* (Econometrica '44)
- Read also
    - *Econometric Causality* (International Statistical Review, 2008)
    - ***Causal Analysis After Haavelmo***, J. Heckman and R. Pinto (Econometric Theory, 2015)
    - *Haavelmo's contributions to simultaneous-equations estimation*, J.S. Chipman (Econometric Theory, 2015)
    - ***Statistical modeling of monerary policy and its effects***, C.A. Sims (Nobel Prize Lecture, December 2011)
    - *A semantic interpretation of Haavelmo's structure of econometrics*, G. Davis (Economics and Philosophy, 2000)

In his words...

1. TRYGVE HAAVELMO'S CAUSALITY

Trygve Haavelmo made fundamental contributions to understanding the formulation and identification of causal models. In two seminal papers (1943, 1944), he formalized the distinction between correlation and causation,[1] laid the foundation for counterfactual policy analysis and distinguished the concept of "fixing" from the statistical operation of conditioning—a central tenet of structural econometrics. He developed an empirically operational version of Marshall's notion of *ceteris paribus* (1890), which is a central notion of economic theory, even though Haavelmo never explicitly used that terminology.

In Haavelmo's framework, the causal effects of inputs on outputs are determined by the impacts of *hypothetical* manipulations of inputs on outputs which he distinguishes from correlations between inputs and outputs in observational data. The causal effect of an input is defined using a hypothetical model that abstracts from the empirical data generating process by making hypothetical variation in inputs that are independent of all other determinants of outputs. As a consequence, Haavelmo's notion of causality relies on a thought experiment in which the model that governs the observed data is extended to allow for independent manipulation of inputs, irrespective of whether or not they vary independently in the data.

Haavelmo formalized Frisch's notion that "causality is in the mind."[2] Causal effects are not empirical statements or descriptions of actual worlds, but descriptions of hypothetical worlds obtained by varying—hypothetically—the inputs determining outcomes. Causal relationships are often suggested by observed phenomena, but they are abstractions from it. [3]

HAAVELMO 44

Mathematics and Economics

One of the most characteristic features of modern economic theory is the extensive use of symbols, formulae, equations, and other mathematical notions. Modern articles and books on economics are "full of mathematics." Many economists consider "mathematical economics" as a separate branch of economics. The question suggests itself as to what the difference is between "mathematical economics" and "mathematics." Does a system of equations, say, become less mathematical and more economic in character just by calling x "consumption," y "price," etc.? There are certainly many examples of studies to be found that do not go very much further than this, as far as economic significance is concerned. But they hardly deserve the ranking of contributions to economics. What makes a piece of mathematical economics not only mathematics but also economics is, I believe, this: When we set up a

Mathematics and Economics: Experiments

mathematics but also economics is, I believe, this: When we set up a system of theoretical relationships and use economic names for the otherwise purely theoretical variables involved, we have in mind some actual *experiment*, or some *design of an experiment*, which we could at least imagine arranging, in order to measure those quantities in real economic life that we think might obey the laws imposed on their theoretical namesakes. For example, in the theory of choice we introduce the notion of indifference surfaces, to show how an individual, at given prices, would distribute his fixed income over the various commodities. This sounds like "economics" but is actually only a formal mathematical scheme, until we add *a design of experiments* that would indicate, first, what real phenomena are to be identified with the theo-

Mathematics and Economics: Experiments

retical prices, quantities, and income; second, what is to be meant by an "individual"; and, third, how we should arrange to observe the individual actually making his choice.

There are many indications that economists nearly always have some such design of ideal experiments in the back of their minds when they build their theoretical models. For instance, there is hardly an economist who feels really happy about identifying current series of "national income," "consumption," etc., with the variables by these names in his theories. Or, conversely, he would often find it too complicated or perhaps even uninteresting to try to build models such that the observations he would like to identify with the corresponding theoretical variables would correspond to those actually given by current economic statistics. In the verbal description of his model, "in economic terms," the economist usually suggests, explicitly or implicitly, some type of experiments or controlled measurements designed to obtain the real variables for which he thinks that his model would hold. That is, he has in mind some "true" variables that he would like to measure. The data he actually obtains are, first of all, nearly always blurred by some plain errors of measurements, that is, by certain extra "facts" which he did not intend to "explain" by his theory. Secondly, and that is still more important, the economist is usually a rather passive observer with respect to important economic phenomena; he usually does not control the actual collection of economic statistics. He is not in a position to enforce the prescriptions of his own designs of ideal experiments.

One could perhaps also characterize the difference between the "true" and the "observational" variables in the following way. The "true" variables are variables such that, if their behavior should contradict a theory, the theory would be rejected as false; while "observational" variables, when contradicting the theory, leave the possibility that we might be trying out the theory on facts for which the theory was not meant to hold, the confusion being caused by the use of the same names for quantities that are actually different.

In order to test a theory against facts, or to use it for predictions, either the statistical observations available have to be "corrected," or the theory itself has to be adjusted, so as to make the facts we consider the "true" variables relevant to the theory, as described above. To use a mechanical illustration, suppose we should like to verify the law of falling bodies (in vacuum), and suppose our measurements for that purpose consisted of a series of observations of a stone (say) dropped through the air from various levels above the ground. To use such data we should at least have to calculate the extra effect of the air resistance and extract this element from the data. Or, what amounts to the same, we should have to expand the simple theory of bodies falling in vacuum, to allow for the air resistance (and probably many other factors). A physicist would dismiss these measurements as absurd for such a purpose because he can easily do much better. The economist, on the other hand, often has to be satisfied with rough and biased measurements.

Models, Hypothesis and Facts

Let $x_1', x_2', \cdots, x_n'$, be n real variables, and let $(x_1', x_2', \cdots, x_n')$, or, for short, (x'), denote any particular set of values of these variables. Any such set may be represented by a point in n-dimensional Cartesian space. Let S be the set of all such points, and let "A" be a system of rules or operations which defines a subset S_A of S. (S_A might, for example, be a certain n-dimensional surface.) The rules "A" ascribe to each point (x') a property, viz., the property of belonging to S_A or not belonging to S_A. If we allow the n variables x' to vary only under the condition that (x') must belong to S_A, this forms a theoretical model for what the variables x' can do.

Similarly, consider n time functions $x_1'(t), x_2'(t), \cdots, x_n'(t)$. Let F be the set of all possible systems of n time functions, and let "B" be a system of rules or operations that defines a subclass F_B of F. Any system of n time functions will then have the property of either belonging to F_B or not belonging to F_B. The system of rules "B" defines a model with respect to n time series.

Models, Hypothesis and Facts

Thus, a theoretical model may be said to be simply a restriction upon the joint variations of a system of variable quantities (or, more generally, "objects") which otherwise might have any value or property. More generally, the restrictions imposed might not absolutely exclude any value of the quantities considered; it might merely give different *weights* (or probabilities) to the various sets of possible values of the variable quantities. The model in question would then usually be characterized by the fact that it defines certain restricted subsets of the set of all possible values of the quantities, such that these subsets have nearly all of the total weight.

Models, Hypothesis and Facts

A theoretical model in this sense is, as it stands, void of any practical meaning or interest. And this situation is, as we have previously explained, not changed by merely introducing "economic names" for the variable quantities or objects involved. The model attains economic meaning only after a corresponding system of quantities or objects in real economic life has been chosen or described, in order to be *identified with those in* the model. That is, the model will have an economic meaning only when associated with a design of actual experiments that describes—and indicates how to measure—a system of "true" variables (or objects) $x_1, x_2, \cdots, x_n$ that are to be identified with the corresponding variables in the theory.

HAAVELMO 43

In the paper where he initiated simultaneous equations modeling, he showed how an hypothesized joint distribution for disturbance terms is transformed by the model into a distribution for the observed data, and went on to show how this allowed likelihood-based methods for estimating parameters. (Sims)

The Statistical Implications of a System of Simultaneous Equations

This is, perhaps, generally realized among econometricians. But they frequently fail to consider, in full, the *statistical* implications of assuming a system of such stochastical equations to be *simultaneously fulfilled* by the data. More specifically, if one assumes that the economic variables considered satisfy, simultaneously, several stochastic relations, it is usually *not* a satisfactory method to try to determine each of the equations separately from the data, without regard to the restrictions which the *other* equations might impose upon the same variables. That this is so is almost self-evident, for in order to prescribe a meaningful method of fitting an equation to the data, it is necessary to define the stochastical properties of *all* the variables involved (e.g., that some of them are given time series, or remain constant, etc.). Otherwise, we shall not know the meaning of the statistical results obtained. Furthermore, the stochastical properties ascribed to the variables in *one* of the equations should, naturally, not contradict those that are implied by the *other* equations. For example, suppose

Section 2

- Present what is (by now) a standard simultaneity problem. Let

$$\begin{aligned} Y &= aX + \epsilon_1, \qquad \epsilon_1 \sim \mathcal{N}\left(0, \sigma_1^2\right) \\ &\Rightarrow \mathbb{E}\left[Y|X\right] = aX \end{aligned}$$

- but

$$X = bY + \epsilon_2, \qquad \epsilon_2 \sim \mathcal{N}\left(0, \sigma_2^2\right)$$

so

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right)$$

and

$$\mathbb{E}\left[Y|X\right] = 0 + \frac{\sigma_Y}{\sigma_X}\rho_{XY}\left(X - 0\right) = \frac{b\sigma_1^2 + a\sigma_2^2}{b^2\sigma_1^2 + \sigma_2^2}X$$

In a system of equations, involving certain specified (but not observable) stochastical variables (ϵ_1, ϵ_2)*, the observable variables involved* (X, Y) *may be considered as **transformations of the specified stochastical ones**. Therefore, the specification of the distribution of these theoretical variables permits us, usually, to calculate the joint distribution of the observable variables, or certain properties of this distribution. This joint distribution should be studied to clarify the stochastical relationship, which the equation system implies with respect to the observable variables. [...]*
*And then, to avoid inconsistencies [...], all formulae for estimating the parameters involved should be derived on the basis of this **joint probability law** of all the observable variables involved in the system.*

- This is the essentially the definition of structural modeling, made general by **Hurwicz (1962)** *On the Structural Form of Interdependent Systems*

Sections 3 and 4

- Formulate a 3-equations dynamic macroeconomic model show the pitfalls in the approach of many economists (of the era) to ... assume that the observed series for c, i, y satisfy the equations "apart from some small deviations, which vary up and down around zero". Then use the time series to least-squares estimate the parameters of the equations separately from one another

$$\hat{\beta} = \arg\min \sum_t (c_t - \beta y_t)^2 \,, \hat{\kappa} = \arg\min \sum_t (i_t - \kappa c_t)^2$$

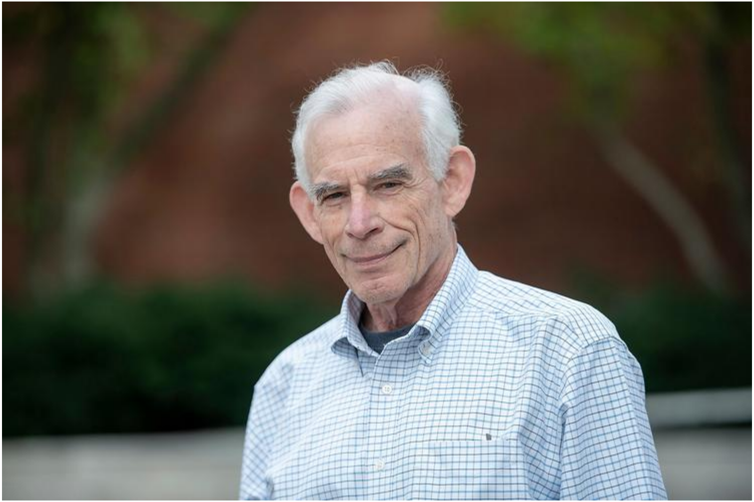
- More on the macroeconomic model with the critiques in the Sims' lecture

Sections 3 and 4

1. *The notion that one can operate with some vague idea about "small errors" without introducing the concepts of stochastical variables and probability distributions, is, I think, based upon an illusion*

2. *Think that the statistical specification job is done when they have stated that: For each point of time, t, [shocks in either equation] are a random variable with a certain probability distribution.*

   - Specifying only the marginal says nothing, and the usual assumption of shocks $\{\epsilon_{c,t}, \epsilon_{i,t}\}_{t\geq 0}$ $\epsilon_{j,s} \perp \epsilon_{k,r}$ $\forall (j,s) \neq (k,r)$ invalidates the OLS procedure.



2011 Nobel Prize Lecture

https://www.nobelprize.org/uploads/2018/06/sims_lecture.pdf

- Word count
  - economic- 52
  - **Haavelmo 41**
  - parameter 40
  - Keynes- (includes Keynesian) 32
  - VAR 22
  - Bayes- (includes Bayesian) 17
  - Friedman 7
  - Sargent 6

Chris Sims Nobel Prize Lecture

- Contextualize 43-44 contributions in the debate (with Keynes) spurred from Tinbergen 39
  - *Keynes irritated reaction to the tedium of grappling with the many numbers and equations [...] finds counterparts to this day in the reaction of some economic theorists to careful, large-scale probability modeling of data. Haavelmo's ideas constituted a research agenda that to this day attracts many of the best economists...*
- Two critiques: *i*) frequentist inference and *ii*) unclear treatment of policy interventions. $\Rightarrow$ Bayesian VARs
  - *The problem of econometric modeling for policy advice is to use the historically estimated joint distribution of policy behavior and economic outcomes to construct accurate probability distributions for outcomes conditional on contemplated policy actions not yet taken.*

Tinbergen-Keynes-Haavelmo

- Tinbergen 39: First multiple-equation, statistically estimated economic time series model.
- Keynes: Can a statistical model like this ever be a framework for testing a theory?
- Haavelmo: Even if they [economic models] are used to make a forecast that is a single number, we understand that the forecast will not be exactly correct
  - We should not give up hope of testing models once we accept that their predictions have error terms [Keynes implication]
  - Can test and compare models, provided that they include a characterization of the nature of their errors
    - Write models as probability distributions for the observed data (43-44 contributions)

Two weaknesses that [...] partially discredited the research program

- Frequentist (Neyman-Pearson) hypothesis testing
  - Analyst cannot assign probability distributions to parameters
  - Limits the possibility to
    1. Studying real-time decision-making under uncertainty (need to assessthe likelihood of various parameter values)
    2. Combine information from model likelihood with beliefs of experts/policymakers
  - Limitations would have been overcome had the literature recognized the value of a Bayesian perspective on inference.
  - *When Haavelmo's ideas were scaled up to apply to models of the size needed for serious macroeconomic policy analysis, the attempt to scale up the hypothesis-testing theory of inference simply did not work in practice.*
  - So, not really his fault.

Two weaknesses that [...] partially discredited the research program

- Failure to confront the conceptual difficulties in modeling policy decisions as both
  - Part of the economic model (also described by probability distributions)
  - Something we can consider changing (and predict how other stuff responds)
- Not even a Lucas critique (any change in policy will systematically alter the structure of econometric models), a policy behavior equation is lacking at all!
- Haavelmo '43 presents a 3 equations model
  1. Consumption equation $c_t = \alpha y_t + \beta + \epsilon_t$
  2. Investment equation $i_t = \kappa c_t + \nu_t$
  3. Accounting identity $y_t = i_t + c_t$

The incriminated part

After showing (Sec 3) that none of the equations 1-3 have no practical significance (in isolation) for prediction purposes (and that specifying the marginal distribution of the structural shocks ϵ isn't enough either), Haavelmo '43 goes on to say...

> What *is* then the significance of the theoretical equations obtained by omitting the error terms in (2.5) and (2.6)? To see that, let us consider, not a problem of passive predictions, but a problem of government planning.
>
> Assume that the Government decides, through public spending, taxation, etc., to keep income, r_t, at a given level, and that consumption u_t and private investment v_t continue to be given by (2.5) and (2.6), the only change in the system being that, instead of (2.7), we now have
>
> $$r_t = u_t + v_t + g_t, \tag{2.7'}$$
>
> where g_t is Government expenditure, so adjusted as to keep r constant, whatever be u and v, as given by (2.5) and (2.6). (2.7′) then does not impose any new restriction upon u and v, beyond that which is ex-

The incriminated part

pressed by (2.5)–(2.6). Then, from (2.5) and (2.6) it is readily seen that

$$E(u_t | r_t) = \alpha r_t + \beta, \tag{4.5}$$

$$E(v_t | u_t - u_{t-1}) = \kappa(u_t - u_{t-1}). \tag{4.6}$$

That is, to predict consumption u_t and private investment v_t under the Government policy expressed by (2.7′) we may use the "theoretical" equations obtained from (2.5) and (2.6) by omitting the error terms x_t and y_t. This is only natural, because now the Government is, in fact, performing "experiments" of the type we had in mind when constructing each of the two equations (2.5) and (2.6).

University Institute of Economics
Oslo

Sims' Conclusion

The confrontation between the monetarists and the Keynesian large scale modelers made clear that econometric modeling of macroeconomic data had not delivered on Haavelmo's research program. He had proposed that economic theories should be formulated as probability distributions for the observable data, and that they should be tested against each other on the basis of formal assessments of their statistical fit. This was not happening. The Keynesians argued that the economy was complex, requiring hundreds of equations, large teams of researchers, and years of effort to model it. The monetarists argued that only a few variables were important and that a single regression, plus some charts and historical story-telling, made their point.

Furthermore, neither side in this debate recognized the centrality of incorporating policy behavior itself into the model of the economy. In the exchanges between Albert Ando and Franco Modigliani (1965) on the one hand, and Milton Friedman and David Meiselman on the other, much of the disagreement was over what should be taken as "autonomous" or "exogenous".