Science, prediction and models *

Jean-Marie Dufour †
McGill University

First version: January 1998
This version: February 11, 2008, 12:10pm

* This work was supported by the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and the Fonds de recherche sur la nature et les technologies (Québec), and a Killam Fellowship (Canada Council for the Arts).2

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

Contents

1. Prediction

1.1 The main objective of scientific theories is to produce "predictions" about observable events. This way of viewing theories is sometimes described as **instrumentalism** [Friedman (1953)].

1.2 "Prediction" is defined here in a wide sense: any type of restriction on the results we can expect from an experiment.

1.3 An "experiment" is also defined in a wide sense: it is any process that can yield "observations" whether the experiment is controlled or purely observational.

1.4 A prediction may be:

(a) *conditional*, such as

"every time the event A occurs,
the event B will occur";

(b) or unconditional. such as

"A will occur"

which is equivalent to

"given the conditions of the experiment
(the state of the world),
A will occur".

Unconditional predictions may be interpreted as special types of conditional predictions.

1.5 Predictions are useful because:

(a) they can be used to make decisions;

(b) they can help us to "explain" phenomena.

"Explaining" a phenomenon is equivalent to being able to make predictions about it.

1.6 A prediction should have two main qualities:

(1) precision: it should be informative;

(2) accuracy: it should be compatible with observations.

These two qualities tend to be *antinomic*: very precise predictions are more easily incompatible with observations.

1.7 If a theory does not lead to predictions about observable events, it is

empirically empty,
devoid of empirical meaning,

and, according to some authors,

non-scientific.

1.8 The more informative (precise) the predictions, the more informative the theory.

1.9 There are two basic ways of imposing restrictions on the results of an experiment:

(a) to define impossible (or sure) events (possibilist prediction schemes); deterministic predictions can be viewed as a special class of possibilist predictions;

(b) to assign probabilities to events (probabilist prediction schemes).

1.10 Scientific theories rarely suggest a single prediction scheme (possibilist or probabilist), but a class of such schemes. In this context,

(a) a *model* can be viewed as a set of prediction schemes;

(b) a *possibilist model* is a class of possibilist prediction schemes; deterministic models can be viewed as a special class of possibilistic models;

(c) a *probabilist (or statistical) model* is a class of probabilist prediction schemes;

(d) an *hypothesis* a subset of a model.

Deterministic models

2. Indeterminacy

2.1 Characteristics of possibilist models _ Possibilist models have two important characteristics.

(a) **Indeterminacy** (Hume, Quine) _ Many prediction schemes (models) are usually compatible with a given set of observations. If we assume one model is the "true" one, there is no way in general to be sure about it.

(b) **Logical falsifiability** (Popper) _ In certain cases, it is possible to conclude that a possibilist model is logically incompatible with a given set of observations.

As a result, possibilist models do not survive easily a confrontation with data.

2.2 Characteristics of probabilist models _ Probabilist models have two important characteristics.

(a) **Indeterminacy**

(b) **Non-falsifiability** _ It is generally impossible to conclude that a probabilist model is logically incompatible with a given set of observations.

The theory of statistical hypothesis tests tries to design "reasonable rules" for accepting or rejecting hypotheses (models).

2.3 Holistic principles (Duhem-Quine) _ Models are usually obtained by combining theoretical hypotheses coming from a theory (e.g., economic theory) with auxiliary assumptions (e.g., distributional assumptions). Since making predictions requires both, it is generally possible to distinguish between theoretical hypotheses "of interest and "auxiliary assumptions".

3. Experiments and models

3.1 We consider an experiment $\mathcal{E}$ whose results belong to s set of possible results $\mathcal{Z}$.

3.2 The symbol Z will denote the realized value of the experiment $\mathcal{E}$, while z will denote any possible result (element) in $\mathcal{Z}$.

3.3 It will be convenient to classify the elements of $\mathcal{Z}$ in subsets. We consider a family $\mathcal{A}_{\mathcal{Z}}$ of subsets of $\mathcal{Z}$. The elements of $\mathcal{A}_{\mathcal{Z}}$ are called *events*.

3.4 Let $A \in \mathcal{A}_{\mathcal{Z}}$. If $Z \in A$, we say "the event A has occurred".

3.5 Usually, the class $\mathcal{A}_{\mathcal{Z}}$ satisfies the properties of an algebra or a $\sigma-$algebra.

3.6 A probabilistic prediction scheme is obtained by assigning a probability to each event in $\mathcal{A}_{\mathcal{Z}}$, i.e. by defining a probability measure on $(\mathcal{Z},\ \mathcal{A}_{\mathcal{Z}},)$. One then obtains in this way a probability space $(\mathcal{Z},\ \mathcal{A}_{\mathcal{Z}},\ P)$.

3.7 A probability (or statistical) model is obtained by considering a set $\mathcal{P}$ of possible probability measures on $(\mathcal{Z},\ \mathcal{A}_{\mathcal{Z}})$ .This yields a triplet of the form $(\mathcal{Z},\ \mathcal{A}_{\mathcal{Z}},\ \mathcal{P})$.

References

DUFOUR, J.-M. (2000): “Économétrie, théorie des tests et philosophie des sciences,” in *Présentations de l’Académie des lettres et des sciences humaines*, vol. 53, pp. 166–182. Royal Society of Canada/Société royale du Canada, Ottawa.

——— (2001): “Logique et tests d’hypothèses: réflexions sur les problèmes mal posés en économétrie,” *L’Actualité économique*, 77(2), 171–190.

FRIEDMAN, M. (1953): “The Methodology of Positive Economics,” in *Essays in Positive Economics*, pp. 3–53. University of Chicago Press, Chicago, Illinois.

HUME, D. (1748): *An Inquiry Concerning Human Understanding*. London.

POPPER, K. (1968): *The Logic of Scientific Discovery*. Harper Torchbooks, New York, revised edn.