

CORE DISCUSSION PAPER

2006/82

THE FIRST STAGE IN HENDRY'S REDUCTION THEORY REVISITED

Genaro SUCARRAT¹

September 2006

Abstract

The reduction theory of David F. Hendry provides a comprehensive probabilistic framework for the analysis and classification of the reductions associated with empirical econometric models. However, it is unable to provide an analysis on the same underlying probability space of the first reduction—and hence the subsequent reductions—given a commonplace theory of social reality, namely the joint hypotheses that the course of history is indeterministic, that history does not repeat itself, and that the future depends on the past. As a solution this essay proposes that the elements of the underlying outcome space in Hendry's theory are interpreted as indeterministic worlds made up of historically inherited particulars.

JEL Classification: B40, C50

Keywords: Theory of reduction, DGP, possible worlds, measurement error, probabilistic causality

¹Departamento de Economía, Universidad Carlos III de Madrid (Spain), and CORE and the Department of Economics, Université catholique de Louvain (Belgium). Email: sucarrat@core.ucl.ac.be. Homepage: <http://www.core.ucl.ac.be/~sucarrat/index.html>.

I am indebted to various people for useful questions, comments and suggestions at different stages, including Farooq Akram, Luc Bauwens, Vincent Bodart, Geert Dhaene, David F. Hendry, Sebastien Laurent, Michel Mouchart, Ragnar Nymoen, Fatemeh Shadman, David Veredas, seminar participants at the department of economics in Oslo in September 2005 and seminar participants at the doctoral workshop in Louvain la Neuve in February 2005. The essay grew out of an earlier which was presented as a poster at the 12th EC-squared conference in Louvain-la-Neuve (December 2001). Comments by conference participants and Alois Kneip, who commented on a precursor to this paper, is also gratefully acknowledged. The usual disclaimer about errors and interpretations being my own applies of course also here. The research has in part been financed by The Norwegian State Educational Loan Fund, Lise and Arnfinn Heje's Fund (Norway), and by the European Community's Human Potential Programme under contract HPRN-CT-2002-00232, MICFINMA.

This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the author.

1 Introduction

When Trygve Haavelmo suggested that the n observations in a dataset could "be considered as *one* observation of n variables... following an n -dimensional *joint* probability law" (1944, p. iii), his main aim was to convert more economists to the praxis of evaluating economic theories against empirical economic data using statistical techniques. The deeper question about how the joint n -dimensional probability distribution was related to reality, however, he remained agnostic about. In his own words, the existence of such a joint probability distribution "may be purely hypothetical" (same place, p. iii).¹ Although Haavelmo's ideas had a profound and immediate impact on contemporary economic analysis it nevertheless took until the 1970s and 1980s before a systematic approach to the study of the relation between reality and models thereof in terms of probability concepts developed in econometrics. At the centre of several important contributions during these years, including Florens and Mouchart (1980, 1985), Hendry and Richard (1990), and Florens et al. (1990), was the notion of a "probabilistic reduction", that is, the idea of replacing a complex probabilistic structure with a simpler one through marginalisation and/or conditioning, and led to the development of important econometric concepts like weak exogeneity, strong exogeneity and super exogeneity (see Engle et al. (1983)).

The reduction theory of David F. Hendry (1995, chapter 9), where the term reduction is used in a broader way than originally, provides a comprehensive probabilistic framework for the analysis and classification of the simplifications associated with empirical models.² In Hendry's own words "it seeks to explain the origin of empirical models in terms of reduction operations conducted implicitly on the DGP [data generating process] to induce the relevant empirical model" (1995, p. 344). Starting with the joint probability distribution of the "complete set" of theory "variables relevant to the economy under investigation" (same place, p. 345), the reduction theory distinguishes between twelve reduction operations which ultimately leads to the empirical model. Although Hendry's theory is a powerful and comprehensive framework for the analysis of the relation between models and reality, it is nevertheless unable to provide an analysis on the same probability space of the first stage—and hence of the subsequent stages—of reduction given a commonplace theory of social reality. The commonplace theory consists of the joint hypotheses that,

¹Later Haavelmo quotes Pareto: "Il n'y a pas de proposition qu'on ne puisse certifier vraie sous certaines conditions, à déterminer" (same place, p. 1). This might be taken as an indication on why he was agnostic about the existence or "truthfulness" of such a joint distribution. Differently put, agreeing with Pareto that truth is a complex issue and highly dependent on tests for it, Haavelmo chose to defer the topic rather than engaging into a detailed and possibly futile discussion.

²Chapter 9 in Hendry (1995) is a revised version of Cook and Hendry (1994), which is based on Hendry and Richard (1990).

literally, a) the course of history is indeterministic, b) history does not repeat itself, and c) the future depends on the past (historical inheritance). In philosophical jargon, that the human world is made up of indeterministic, historically supervenient particulars.³ According to Hendry the economic mechanism under study, that is, the joint distribution of the complete set of relevant theory variables defined on the underlying probability space, is an entity that *can* change but does not necessarily do so. According to the commonplace theory of social reality on the other hand, relationships change all the time in a way that is indeterministically (but not stochastically) dependent on the past. Conceptually this is not necessarily incompatible with Hendry's view, but it would nevertheless imply that the economic mechanism is constantly changing. As a consequence, Hendry's theory is unable to provide a probabilistic analysis on the same underlying probability space when relationships change. Moreover, according to Hendry the underlying probability space is transformed—again—when data are collected, so the theory is unable to provide a probabilistic analysis on the same underlying probability space of the relation between the theory and data variables.

In this essay I propose a certain structure on the underlying outcome space with the consequence that the associated probability space remains the same throughout all reductions in Hendry's theory. The proposed structure consists of devising the outcome set as consisting of possible worlds made up of indeterministic and historically inherited particulars, and provides several gains and possibilities of which only a few are explored in this essay. First, the formulation or choosing of theory variables can be seen as a simplification or some sort of "pre-marginalisation" of variables, and can thus be interpreted as the "perspective" from which we study an issue. Second, a probabilistic definition of measurement validity, that is, the absence of measurement error, is enabled. Third, a history based probabilistic definition of indeterministic causality that nests discrete, continuous and "interval" versions of probabilistic causality is proposed. Finally, mathematical expectation conditional on an information set is re-interpreted.

More generally devising the outcome set as consisting of indeterministic worlds made up of historically inherited particulars provides a bridge between econometric (/probabilistic) reduction analysis and metaphysics, that is, the part of philosophy that deals with what there exists and its nature. There is already a voluminous philosophical literature that employs the idea of possible worlds to shed light on various metaphysical issues, and by providing a bridge between these two literatures econometrics can benefit from these insights.

³Throughout the essay I will employ philosophical terms which I explain only briefly. Readers interested in fuller explanations or further reading are referred to (say) Honderich (1995) and Craig (2000).

The rest of this essay is organised into five sections. In the next, section 2, the most relevant parts of Hendry’s reduction theory is detailed. Section 3 describes and motivates the structure of the outcome space that is proposed. Section 4 explains why the probability space remains the same throughout the reductions, details in what sense the formulation of theory variables can be seen as a simplification, and proposes probabilistic definitions of measurement validity, that is, absence of measurement error. Section 5 formulates the proposed definition of causality and re-interprets mathematical expectation conditional on an information set. Finally, section 6 concludes and gives suggestions for further research.

2 The first stage in Hendry’s reduction theory

The purpose of Hendry’s reduction framework is ”to explain the origin of empirical models in terms of reduction operations conducted implicitly on the DGP” (1995, p. 344), and his framework details twelve reductions whose order is not unique.⁴ Since the focus in this essay is on the first stage I concentrate on this in what follows.

The most informative account of the first stage of reduction is given in a single paragraph in chapter 9 of *Dynamic Econometrics* (1995), which is an adaptation of Hendry and Cook (1994). Most of the paragraph is about the concepts and actions involved in the first stage, so it seems useful to reproduce it here almost in its entirety. Note however that I have modified Hendry’s notation in order to retain a consistent notation throughout this essay. Most importantly, random variables and vectors appear in capitals to distinguish them from their realisations, which I denote in small letters later in the essay. The passage is:

”The analysis begins with the complete set of random variables $\{\mathbf{U}_t^*\}$ relevant to the economy under investigation over a time span $t = 1, \dots, T$, where the superscript * denotes a perfectly measured variable $\mathbf{U}^* = (\mathbf{U}_1^*, \dots, \mathbf{U}_T^*)$, defined on the probability space $(\Omega, \mathcal{F}, P) \dots$ The $\{\mathbf{U}_t^*\}$ comprise all the potential variables from the economic mechanism under study which operates at the level of \mathbf{U}^* , and hence the vector \mathbf{U}_t^* comprises details of every economic action of every agent at time t in all the regions of the geographical space relevant to the analysis. However, many of the $\{U_{ti}^*\}$ variables are either unobserved or badly measured, so the term data is not strictly applicable to \mathbf{U}_t^* . The mapping from

⁴The ”important point”, he says, ”is that empirical relationships must arise from these reductions of the DGP” (same place, p. 345).

the economic mechanism to the data-generation process through the measurement system is the first reduction, which can lose a vast amount of information, and introduce inaccuracy but leads to a data-set which is denoted by $\{\mathbf{U}_t\}$. At a conceptual level, all variables $\{U_{ti}^*\}$ are assumed to be measured as $\{U_{ti}\}$ although for some variables, the level of quantification may be low, possibly even an artificial entry of zero. The probability space (Ω, \mathcal{F}, P) is transformed by the measurement process (usually markedly) . . .”—Hendry (1995, p. 345)

Thus the starting point of Hendry’s reduction theory is a set of theory variables denoted \mathbf{U}^* defined on the probability space (Ω, \mathcal{F}, P) , and together \mathbf{U}^* and (Ω, \mathcal{F}, P) constitute the ”economic mechanism”. Furthermore, the actions of collecting and recording the data, that is, the measurement process, produces a dataset \mathbf{U} defined on an altered probability space $(\Omega', \mathcal{F}', P')$. This altered probability space $(\Omega', \mathcal{F}', P')$ together with the data variables \mathbf{U} is called the ”data generating process” (DGP). Schematically the first stage of reduction is summarised in table 1.

3 The outcome set as consisting of possible worlds

If (Ω, \mathcal{F}, P) denotes a probability space with Ω , \mathcal{F} and P being the outcome space, the event space and the probability measure, respectively, then in what follows the elements $\omega \in \Omega$ will be referred to as ”worlds” or ”possible worlds”. The purpose of this section is to formulate and motivate the proposed structure of the worlds ω . The proposed structure serves as some sort of social ontology, that is, a theory of the nature of social reality, and is contained in definition 4 in the last subsection of this section. The preceding subsections provide the details that leads up to definition 4. The first subsection 3.1 presents the idea of a possible world which in philosophy has proved very useful in analysing, discussing and communicating many philosophical ideas and theories, and shows that there is no loss of generality in interpreting the ω as worlds. Then, subsections 3.2 and 3.3 formalise the ideas of contingent particularism and historically inherited particulars, respectively. Finally, subsection 3.4 contains the definition of outcomes sets consisting of indeterministic worlds made up of historically inherited particulars.

3.1 Possible worlds

The idea of a world is normally credited to the German philosopher and mathematician Gottfried Wilhelm Leibniz (1646 - 1716) (Crane 1995).⁵ Intuitively a world contains

⁵Leibniz was religious and originally he used the idea to argue that the world is perfect because among all the possible worlds God must have chosen the most perfect one, an idea that was ridiculed by Voltaire

everything in the past, everything in the present and everything in the future, or in Leibniz' own words "the entire sequence and the entire collection of all existing things" (*Theodicy*, par. 8, G VI 107. Quoted in Parkinson 1995, p. 213). In contemporary philosophy the notion is often associated with David Lewis (1941-2001), who describes worlds as consisting of

"the planet Earth, the solar system, the entire Milky Way, the remote galaxies we see through telescopes. . . Anything at any distance at all is to be included. Likewise the world is inclusive in time. No long-gone ancient Romans, no long-gone pterodactyls, no long-gone primordial clouds of plasma are too far in the past, nor are the dead dark stars too far in the future, to be part of this same world"—Lewis (1986b, p. 1)

Where Lewis differ from Leibniz is with respect to how things are connected and with respect to the existence of other worlds. Whereas Leibniz was a determinist and believed in a single world, Lewis was a non-determinist and believed in the rather unusual thesis that non-actual, possible worlds exist objectively and independent of thought because "philosophy [his own?] goes more easily" if we believe so (1986b, p. vii).⁶ Although I take side with Lewis in the determinism *vs.* indeterminism debate my view differs most certainly from Lewis' regarding the existence of non-actual worlds, since I only see them as useful mind-constructs not existing independent of thought.⁷

But do we really need the whole world for the purpose of econometric reduction analysis? Spatially, yes, if we want to ensure a complete analysis, but it is not necessary to be all-including backwards and forward in time. Differently put, the worlds must contain everything between a start point and an end point, but the portions outside this interval are not really necessary although including them changes little. So henceforth I will devise a world ω as a continuous time process $\{w(t) : t \in [0, \infty)\}$ of worldly states-of-affairs $w(t)$, where $[0, \infty)$ is contained in the set of real numbers \mathbb{R} . The number 0 denotes an arbitrary starting point, say, yesterday at midnight or four million years ago, and is not restrictive. However, bounding worlds temporally backwards in time entails an implicit conditioning on the realised history preceding 0. Backwards bounding thus means probabilities acquire

in his play *Candide* (Crane 1995). In today's philosophical usage however the term usually carries no religious connotation.

⁶The whole book is a defence of this thesis but see in particular pp. vii-ix and pp. 133-135. For a brief and amusing summary of other philosophers' reactions to Lewis's thesis, see Hawthorn (1995, footnote 24 pp. 23-24).

⁷For further philosophical issues and references regarding the idea of a possible world useful starting points are Forbes (1995) and Moravcsik (1995). For an alternative but related use of the idea of a possible world by an economist, see Kluve (2004).

an interpretation of special interest, but apart from this the only function bounding serves is to simplify the exposition.

Interpreting ω as worlds retains the intuitive use of probability algebra. For example, if we want to say that $A \in \mathcal{F}$ denotes the event that (say) 10% of the labour force of an economy is unemployed at t , then the only change in interpreting the ω as a world is that A now denotes the set of all worlds in which 10% of the labour force of a certain economy is unemployed at t . More formally, $A = \{\omega : 10\% \text{ unemployed at } t\}$. If the worlds are bounded backwards, then the interpretation becomes that A denotes the set of all worlds in which 10% of an economy is unemployed at t given the history of the world up to $t = 0$. Another common practice is to interpret the outcome set Ω as a set of possible "states-of-affairs" or "facts". In possible worlds terminology a state-of-affair or fact at t is now the set of all worlds in which a certain state-of-affairs or fact attains at t . Finally, the possible worlds interpretation also accommodates "interval" events. With respect to the unemployment example, the event A now becomes the set of worlds in which 10% of the labour force of an economy is registered as unemployed over the time interval, say, $[t_0, t_1]$.

3.2 Contingent particularism

"I am inclined", in the words of Geoffrey Hawthorn, "to the view that the human world consists of contingent particulars" (1995, p. 10). Contingency refers to the thesis that social events are not connected in a deterministic manner, a question that has occupied philosophers for thousands of years. There are at least two philosophical literatures of relevance for this issue. The first is concerned with whether human being is endowed with a so-called "free will" and if so what kind of free will. The second literature is the so-called "philosophy of mind" literature and starts from two seemingly contradictory views. On the one hand that human being presumably is made up of a finite number of indivisible objects, usually referred to as particles, and on the other hand that human being is capable of a presumably infinite number of mental states (imagination, thought, etc.).⁸ Depending on one's views on free will and on the relationship between mind and matter, a variety of possible views on how social events are connected is possible. Since I am unlikely to convince the reader of my belief in the indeterminism thesis unless she or he is already

⁸An entry on "free will" is contained in virtually any philosophy or metaphysics dictionary, see for example Honderich (1995) or Kim and Sosa (1995), and usually contain further reading. A very accessible introduction to these issues, which is based on the author's BBC lectures, is Searle (1991). Useful introductions to the philosophy of mind are Kim (1996) and Heil (1998), the second being more advanced than the first. A good text on the relation between mind and recent biological currents is Ruse (1988). Texts that consider themselves to specifically address issues of social ontology are Ruben (1985) and Pettit (1993). A useful introduction to metaphysics as it is often conceived, a form of category theory, is Loux (1998).

a believer I merely state the thesis as some sort of axiom that I start from. Formally, with respect to the probability space (Ω, \mathcal{F}, P) , indeterminism is simply characterised by Ω containing more than one element, that is, part a) in definition 4.

The meaning of the philosophical idea of a "particular" is best understood when contrasted with its opposite, a "universal". In brief, something is said to be of particular nature if there exists only one of its kind, whereas something is said to be of universal nature if it is one out of several of its kind or type. Another way to put it is that a particular refers to the unique and non-repeatable, whereas a universal refers to the repeatable. In the current context particularism is the thesis that, literally, history does not repeat itself (no two points in time are exactly equal in all respects).⁹ Formally this may be stated as follows.

Definition 1. Worldly particularism. A world $\omega = \{w(t) : t \in [0, \infty)\} \in \Omega$ is said to be made up of particulars if for all pairs $t, t' \in [0, \infty)$ such that $t \neq t'$ and $w(t), w(t') \in \omega$, then $w(t) \neq w(t')$.

3.3 Historically inherited particulars

A further thesis I start from is that the current and the future depends on and inherit the characteristics of the past. Differently put, every turn history takes contributes in one or another way to the characteristics of the worldly state-of-affairs of the future. This thesis I shall call "historical inheritance", but before providing a formal formulation of this property we need the idea of a worldly state-of-affairs process up to t .

Definition 2. Worldly state-of-affairs process. The process $\omega_t = \{w(a) : a < t, t \in (0, \infty)\} \subsetneq \omega$ is said to be a worldly states-of-affairs process up to but not including t .

So intuitively ω_t is a history up to t and note that the number 0 is not included in the interval $(0, \infty)$ in order to ensure that ω_t is non-empty. We can now define historical inheritance.

⁹A further interpretation of the thesis that the human world is made up of particulars is that, literally, people differ from each other: No two persons are equal in all respects at any point in time. In the current context, however, we only need the first interpretation.

Definition 3. Historical inheritance. The outcome space Ω is said to consist of worlds made up of historically inherited particulars if:

- a) All $\omega \in \Omega$ are made up of particulars.
- b) For all pairs of unequal worlds $\omega^1, \omega^2 \in \Omega$, that is, $\omega^1 \neq \omega^2$: If $\omega_t^1 \neq \omega_t^2$, then $w^1(t') \neq w^2(t')$ for all $t' \in [t, \infty)$, where $w^1(t') \in \omega^1$ and $w^2(t') \in \omega^2$.

In words, if two worlds contains the same history up to t (but not at t), then the two worlds differ from each other in at least one respect at every point in the future, that is, from t and onwards.

3.4 Outcome sets consisting of indeterministic worlds made up of historically inherited particulars

The proposed structure of the worlds ω is contained in definition 4. The definition summarises the ideas of this section and provides the starting point for what follows.

Definition 4. Outcome set consisting of indeterministic worlds made up of historically inherited particulars. Let (Ω, \mathcal{F}, P) be a probability space and let each $\omega \in \Omega$ be equal to a non-stochastic continuous time process $\{w(t) : t \in [0, \infty)\}$ with $[0, \infty) \subset \mathbb{R}$. The outcome space Ω is said to consist of possible worlds made up of indeterministic and historically inherited particulars if:

- a) There exists more than one element in Ω (indeterminism).
- b) For each $\omega \in \Omega$: For all pairs $t, t' \in [0, \infty)$ such that $t \neq t'$ and $w(t), w(t') \in \omega$, then $w(t) \neq w(t')$ (particularism).
- c) For each pair of unequal worlds $\omega^1, \omega^2 \in \Omega$, that is, $\omega^1 \neq \omega^2$: If $\omega_t^1 \neq \omega_t^2$ then $w^1(t') \neq w^2(t')$ for all $t' \in [t, \infty)$, where $w^1(t') \in \omega^1$ and $w^2(t') \in \omega^2$ (historical inheritance).

The first property a) essentially states that the course of history is indeterministic. If Ω contained only a single world, then this would imply that no other worlds are possible and therefore that the course of history is deterministic. The second property b) makes use of the notion "states-of-affairs" at t which is denoted $w(t)$, and essentially states that history does not repeat itself. The third and final property c) imposes a certain structure on the history-does-not-repeat itself property. Specifically, it ensures that future properties are shaped by the past, thus the terminology "historical inheritance".

4 The first stage in Hendry’s reduction theory revisited

The probability space of definition 4 in subsection 3.4—denoted (Ω, \mathcal{F}, P) —and a set of theory variables \mathbf{U}^* provide the starting point of this section. Together they constitute the ”revised” economic mechanism. Subsection 4.1 explains why (Ω, \mathcal{F}, P) does not change although social relationships change, whereas subsection 4.2 outlines an interpretation of the formulation of theory variables that is of practical use in empirical econometrics. Finally, subsection 4.3 proposes a formal definition of measurement validity, that is, the absence of measurement error, which is enabled by the property that (Ω, \mathcal{F}, P) does not change due to the measurement process.

4.1 An unchanging probability space

A useful distinction is that between social relationships on the one hand and representations thereof on the other. The economic mechanism, that is, \mathbf{U}^* together with (Ω, \mathcal{F}, P) , is a probabilistic representation of social relationships. In other words, the economic mechanism is a picture of the landscape it depicts, not the landscape itself. When the landscape change, so does the picture of it. This is Hendry’s view in brief. The commonplace theory of social reality, which holds that social reality and the associated relationships are changing all the time since they would hinge on their exact historical and geographical location, would therefore imply that the economic mechanism is changing all the time. An unchanging probability space—even when social relationships are changing all the time—is entailed by interpreting the outcome set as consisting of indeterministic worlds made up of historically inherited particulars.

To see this consider the example of a two period series of 0-1 variables $\{X_1, X_2\}$, that is, for $t = 1, 2$ each X_t can take on the values 0 and 1. If $p_{x_1x_2}$ denotes the joint probability of $X_1 = x_1$ and $X_2 = x_2$, then the set of joint probabilities $\{p_{00}, p_{01}, p_{10}, p_{11}\}$ is denoted \mathbf{p} . In other words, \mathbf{p} is a probabilistic representation of the social relationships between X_1 and X_2 . If the social relationships change, then the corresponding probabilistic representation changes from \mathbf{p} to \mathbf{p}' . But must the underlying probability space change? Not if Ω is interpreted as consisting of indeterministic worlds made up of historically inherited particulars. Because then \mathbf{p} and \mathbf{p}' can be treated as conditional on (say) a third random variable that governs when \mathbf{p} and \mathbf{p}' hold. Specifically, let Y denote such a third random variable which is equal to 0 whenever \mathbf{p} holds, that is, $\{\omega : Y(\omega) = 0\}$ is the set of all worlds in which \mathbf{p} holds, and which is equal to 1 whenever \mathbf{p}' holds, that is, $\{\omega : Y(\omega) = 1\}$ is the set of all worlds in which \mathbf{p}' holds. Intuitively, you may interpret $Y = 0$ as (say) the absence of nuclear world wars for $t = 1, 2$, and $Y = 1$ as their presence

for $t = 1, 2$. This gives $\mathbf{p} = \{p_{000}, p_{010}, p_{100}, p_{110}\}$ and $\mathbf{p}' = \{p_{001}, p_{011}, p_{101}, p_{111}\}$, and provides four conditions for $\mathbf{p} = \mathbf{p}'$ to hold: $p_{000} = p_{001}, p_{010} = p_{011}, p_{100} = p_{101}$ and $p_{110} = p_{111}$. In other words, the presence of an unchanging joint density of the theory variables can be treated as conditional on the worlds in which it does not change, and hence the existence of an unchanging economic mechanism constitutes a reduction.

4.2 Formulation of theoretical variables as a reduction

Normative analysis is about how things should be, it is said, whereas positive analysis is value-independent and "objective" investigation of how things are. But is positive analysis entirely objective? Do we not, in any investigation, choose which questions to address, which portions of social reality to study, and which categorical schemes, concepts, techniques and language to employ? The idea that these choices are non-objective in some sense is old and not controversial. Examples of economists who held this view are Max Weber (1994), Joseph Schumpeter (1949) and Gunnar Myrdal (1953, pp. vii-viii; 1969). Since a world contains everything and since the outcome set contains all the possible worlds, the formulation of theoretical variables defined on the probability space can be seen as reflecting some of these choices. In particular, the formulation of theoretical variables can be seen to reflect which portions of reality that are studied as opposed to others, that is, as some sort of "pre-marginalisation" of the variables that are not studied. (The term "pre-marginalisation" is due to the fact that marginalisation of the unimportant—according to some criteria—variables of the set that is considered takes place later in the reduction process.) Differently put, the formulation of theoretical variables can be seen as the "conceptual lenses" we view reality with. For example, in delineating and defining theoretical price and theoretical quantity, then other aspects of the transaction process are not included in the analysis. This is clearly an abstraction, since an anthropologist or an institutional economist might be interested in whether the parties engaged in any form of negotiation, whether there were implicit power-relations governing the transaction process, or what the means of transactions were. All this and many other aspects of the transaction are excluded from the analysis when the only theoretical variables delineated are price and quantity. With the modified probability space the selection of which portions of reality to analyse and the way they are depicted in terms of variables can be treated as a simplification.

4.3 A probabilistic definition of measurement validity

In the methodological literature of the social sciences, discussions of measurement error are often couched in terms of theoretical or nominal or concept definition *vs.* measure or indicator or operational definition—see for example de Vaus (2001, pp. 24-33), Punch (1998, pp. 47-48) and Crano and Brewer (2002, pp. 5-12). That is, to what extent a measure (say, the number of people receiving unemployment benefits) is capable of providing information about a theoretical definition (say, the number of unemployed). An operational definition that satisfactorily provides the information sought is thus said to be measurement valid or concept valid.

To see that the collection of data variables does not alter the underlying probability space, recall that any realisation of the data variable \mathbf{U} corresponds to the worlds in which the data were collected or could have been collected. For example, for any realisation \mathbf{u}_t of \mathbf{U}_t there is an associated set of possible worlds $\{\omega : \mathbf{U}_t(\omega) = \mathbf{u}_t\}$ in which these data realisations can be obtained. Also, if we would like to restrict ourselves to the worlds enabled by history, then we can restrict ourselves to the intersection of $\{\omega : \mathbf{U}_t(\omega) = \mathbf{u}_t\}$ and the set of possible worlds enabled by the course of history preceding t .

To repeat, random variables are denoted in capitals and their realisation in small letters. For example, a realisation of the theoretical vector of variables \mathbf{U}^* is denoted $\mathbf{u}^* = (\mathbf{u}_1^*, \mathbf{u}_2^*, \dots, \mathbf{u}_t^*, \dots, \mathbf{u}_T^*)$, with $\mathbf{u}_t^* = (u_{t1}^*, u_{t2}^*, \dots, u_{ti}^*, \dots, u_{tI(t)}^*)$ for each t , where the symbolism $I(t)$ means the number of theoretical variables can vary with t . Similarly, a realisation of the vector of data variables \mathbf{U} is denoted $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_t, \dots, \mathbf{u}_T)$, with $\mathbf{u}_t = (u_{t1}, u_{t2}, \dots, u_{tj}, \dots, u_{tJ(t)})$ for each t , where the symbolism $J(t)$ means the number of data variables can vary with t . $J(t)$ may of course differ from $I(t)$. Ideally a definition of measurement validity of \mathbf{U}^* should be sequential and formulated for a sequence of pairs $(\mathbf{U}_1^*, \mathbf{U}_1), (\mathbf{U}_2^*, \mathbf{U}_2), \dots, (\mathbf{U}_t^*, \mathbf{U}_t), \dots, (\mathbf{U}_T^*, \mathbf{U}_T)$, where at each t one may (or may not) condition on history and/or on data realisations preceding t . However, such a definition complicates notation considerably so I only provide the definition for a generic t only, $(\mathbf{U}_t^*, \mathbf{U}_t)$, since the extension to $t = 1, 2, \dots, T$ is straightforward. Now, recall the definition of a measurable variable.

Definition 5. Measurable variable. Let (Ω, \mathcal{F}) and $(\Omega^*, \mathcal{G}^*)$ denote two measurable spaces, that is, \mathcal{F} and \mathcal{G}^* are σ -fields on Ω and Ω^* , respectively, and denote the elements of \mathcal{F} and \mathcal{G}^* for F and G^* , respectively. A function $f : \Omega \rightarrow \Omega^*$ is said to be \mathcal{F} -measurable if for all $G^* \in \mathcal{G}^*$ we have $\{\omega : f(\omega) \in G^*\} \in \mathcal{F}$.

In the case where Ω^* is Euclidean space then f is a random vector. For notational convenience

nience I will use the symbolism $f : (\Omega, \mathcal{F}) \longrightarrow (\Omega^*, \mathcal{G}^*)$ to mean that f is a \mathcal{F} -measurable function from Ω to Ω^* , with \mathcal{F} and \mathcal{G}^* being the associated σ -fields. Now, consider the two measurable variables

$$\mathbf{U}_t^* : (\Omega, \mathcal{F}) \longrightarrow (\mathbf{X}_t^*, \mathcal{G}_t^*) \text{ and } \mathbf{U}_t : (\Omega, \mathcal{F}) \longrightarrow (\mathbf{X}_t, \mathcal{G}_t),$$

where $\mathbf{X}_t^* = X_{t1}^* \times X_{t2}^* \times \cdots \times X_{tI(t)}^*$ and $\mathbf{X}_t = X_{t1} \times X_{t2} \times \cdots \times X_{tJ(t)}$, and think of the first as the theory variable and the second as the data variable. The elements of \mathcal{F} , \mathcal{G}_t^* and \mathcal{G}_t will be referred to as worldly events at t , theory events at t and data events at t , respectively. Measurement validity of the data event $G_t \in \mathcal{G}_t$ with respect to the theoretical event $G_t^* \in \mathcal{G}_t^*$ can now be defined in terms of equality between the worldly events $\{\omega : \mathbf{U}_t^*(\omega) \in G_t^*\} \in \mathcal{F}$ and $\{\omega : \mathbf{U}_t(\omega) \in G_t\} \in \mathcal{F}$. In words, to what extent the set of possible worlds associated with a certain data realisation equals the set of worlds associated with the theory event it purports to measure. Generalised the idea can be summarised in the following definition.

Definition 6. Measurement validity of data events. A data event $G_t \in \mathcal{G}_t$ is said to be:

- a) *measurement valid* with respect to a theory event $G_t^* \in \mathcal{G}_t^*$ if $\{\omega : \mathbf{U}_t(\omega) \in G_t\} = \{\omega : \mathbf{U}_t^*(\omega) \in G_t^*\}$.
- b) *measurement invalid* with respect to a theory event $G_t^* \in \mathcal{G}_t^*$ if $\{\omega : \mathbf{U}_t(\omega) \in G_t\} \cap \{\omega : \mathbf{U}_t^*(\omega) \in G_t^*\} = \emptyset$.
- c) *partially measurement valid* with respect to a theory event $G_t^* \in \mathcal{G}_t^*$ if $\{\omega : \mathbf{U}_t(\omega) \in G_t\} \neq \{\omega : \mathbf{U}_t^*(\omega) \in G_t^*\}$ and $\{\omega : \mathbf{U}_t(\omega) \in G_t\} \cap \{\omega : \mathbf{U}_t^*(\omega) \in G_t^*\} \neq \emptyset$.

For convenience we may say that a data event is measurement valid, invalid or partially valid, respectively, since it is implicitly understood that the validity is with respect to a certain theory event. The extensions to theoretical variables is more or less straightforward, but for convenience I only provide the definition for measurement validity.

Definition 7. Measurement validity of a data variable. A data variable $\mathbf{U}_t : (\Omega, \mathcal{F}) \longrightarrow (\mathbf{X}_t, \mathcal{G}_t)$ is said to be measurement valid if each $G_t \in \mathcal{G}_t$ is measurement valid.

Implicitly the definition thus assumes there is a theory variable $\mathbf{U}_t^* : (\Omega, \mathcal{F}) \longrightarrow (\mathbf{X}_t^*, \mathcal{G}_t^*)$ defined on the probability space (Ω, \mathcal{F}, P) . Finally, a definition of almost sure measurement validity can be formulated.

Definition 8. Almost sure measurement validity of a data variable.

Consider a data variable $\mathbf{U}_t : (\Omega, \mathcal{F}) \longrightarrow (\mathbf{X}_t, \mathcal{G}_t)$ and denote the set containing measurement valid data events for $\mathcal{G}_t^1 = \{G_t \in \mathcal{G}_t : G_t \text{ is measurement valid}\} \subset \mathcal{G}_t$. If $P[\bigcup_{i=1}^{\infty} G_t^1(i)] = 1$ where $\{G_t^1(1), G_t^1(2), \dots, G_t^1(i), \dots\} = \mathcal{G}_t^1$, then \mathbf{U}_t said to be *measurement valid almost surely*.

The modified framework is summarised in table 2.

5 A history based probabilistic definition of indeterministic causality

Discussions over what the appropriate definition of causality is for econometrics enjoys a reasonably long history, for overviews and references see amongst others Geweke (1984), Aigner and Zellner (1988), and Bauwens et al. (forthcoming). Most of the suggested definitions have put more weight on empirical implementability rather than philosophical justification, which is understandable given econometrics' nature. In this sense the definition outlined here distinguishes itself by explicitly giving more weight to philosophical considerations rather than empirical implementation. Indeed, I believe that the principal use of the notion of causality proposed here is conceptual analysis rather than empirical analysis. The main characteristic of the definition is that it conceives causality as having two aspects, historical possibility and causal efficiency, and the section proceeds in four steps. In the first subsection the idea of historical possibility is introduced and discussed, and in the second causal efficiency. Subsection 5.3 brings out the most important similarities and differences between the proposed definition and David Lewis' ideas, whereas the final subsection relates the ideas in this section to a common definition of causality in econometrics, namely mathematical expectation conditional on an information set.

5.1 Historical possibility

The first aspect of causality is in a sense obvious. How can an event C be considered as a cause of another event E if the second event is not even possible given the first? The approach to possibility pursued here is that of historical possibility, that is, the idea that what is possible tomorrow depends crucially on where we stand today. In other words, the course of history up to t determines what is possible at and after t . Before we can define this idea formally we need a definition of history.

Definition 9. History up to t . Let ω_t be a state-of-affairs process up to t . The event $H_t = \{\omega : \omega_t \subsetneq \omega\} \in \mathcal{F}$ is said to be a history up to t .

In words H_t is the set of all possible worlds that contain the state-of-affairs process ω_t and intuitively H_t is exactly what its name suggests, namely history up to t . Now, a possible or historically possible event is defined as follows.

Definition 10. A historically possible event. Let $H_{t_1}, E_{t_2} \in \mathcal{F}$ where $t_1 \leq t_2$ and where H_{t_1} is a history up to t . E_{t_2} is said to be a possible event with respect to the history H_{t_1} if $H_{t_1} \cap E_{t_2} \neq \emptyset$.

In words, the event E_{t_2} at t_2 is said to be historically possible or possible for short if at least one of its worlds is contained in history. Similarly, an event is impossible if $E_{t_2} \cap H_{t_1} = \emptyset$, since H_t by construction contains the set of all possible worlds containing the course of history up to and including t . A consequence of definition 10 is that situations where $E_{t_2} \cap H_{t_1} \neq \emptyset$ and $P(E_{t_2} \cap H_{t_1}) = 0$, that is, that E_{t_2} is possible but probabilistically impossible, are not excluded from the outset. Situations where the effect precedes its cause are on the other hand excluded from the outset by the condition $t_1 \leq t_2$. Another characteristic of the definition is that it allows for events being causal for some t but not necessarily at all t . In particular, the definition allows for so-called "single case" causality with "many case" causality being obtained as a probabilistic reduction. Finally, the definition resolves a problem discussed by Salmon (1993b, 1993a). In his view the definitions of probabilistic causality put forward by Reichenbach (1956), Good (1961, 1962 and 1963) and Suppes (1970) all suffer from the fact "that they attempt to carry out the construction of causal relations on the basis of probabilistic relations among discrete events. . ." (1993b, p. 151). In other words, they fail to take into account the continuous processes that connect events. Salmon (1993a) himself proposed a solution that takes "processes rather than events as basic entities" (same place, p. 155). The current approach follows in the same vein and thus constitutes an alternative to Salmon's approach. Specifically the current approach takes continuous states-of-affairs processes (that is, worlds) as basic entities with the consequence that discrete and continuous accounts—indeed, even "interval" accounts—of causality are reconciled in a neat manner.

5.2 Causal efficiency

Defining possibility in this way means conditional probability suggests itself as a measure of causal efficiency. Heuristically a cause is said to be more efficient than another if the first is more likely to bring about the event in question, and formally we may define this as follows:

Definition 11. Causal efficiency. Let $H_{t_1}^a, H_{t_1}^b \in \mathcal{F}$ denote two different histories, that is, $\omega_{t_1}^a \neq \omega_{t_1}^b$, and consider the event $E_{t_2} \in \mathcal{F}$ where $t_1 \leq t_2$. Further let $P(E_{t_2} \cap H_{t_1}^a), P(E_{t_2} \cap H_{t_1}^b) \neq \emptyset$ and $P(H_{t_1}^a), P(H_{t_1}^b) > 0$:

- a) If $P(E_{t_2}|H_{t_1}^a) > P(E_{t_2}|H_{t_1}^b)$, then $H_{t_1}^a$ is said to be causally more efficient than $H_{t_1}^b$ in bringing about E_{t_2} .
- b) If $P(E_{t_2}|H_{t_1}^a) = P(E_{t_2}|H_{t_1}^b)$, then $H_{t_1}^a$ and $H_{t_1}^b$ are said to be causally equally efficient in bringing about E_{t_2} .

As an example, let $H_{t_1}^a$ and $H_{t_1}^b$ denote two different policy choices, say, increasing the interest rate with 0.5%-point and no-change, respectively, and let E_{t_2} denote the desired policy objective, say, a yearly inflation of 2.5% two years into the future. The conditions $E_{t_2} \cap H_{t_1}^a \neq \emptyset$ and $E_{t_2} \cap H_{t_1}^b \neq \emptyset$ essentially state that the policy choices in question both are capable of bringing about the desired objective E_{t_2} . So if, say, $P(E_{t_2}|H_{t_1}^a) > P(E_{t_2}|H_{t_1}^b)$, then alternative a is more likely to attain E_{t_2} than alternative b .

How does all this relate to the more common idea of an event C_{t_1} being the cause of an effect E_{t_2} , most often expressed in terms of $P(E_{t_2}|C_{t_1})$? The answer lies in the structure of the underlying outcome space. Recall that each element in the outcome space is devised as a continuous time process of states-of-affairs from 0 and onwards. This means $P(E_{t_2}|C_{t_1})$ can be interpreted as the probability of the event E_{t_2} given the event C_{t_1} , and given history up to but not including $t = 0$. This is why bounding worlds backwards produces a particularly interesting interpretation of conditional probabilities. Moreover, it is not necessary for worlds to be unbounded forward for this interpretation to obtain. Worlds may just as well be devised as finite non-stochastic continuous time processes starting at $t = 0$ and ending when $t = T$. Similarly, the probability $P(E)$ where E is an arbitrary event in the event-set can be interpreted as the probability of E coming about *conditional* on history up to but not including $t = 0$.

5.3 David Lewis' ideas compared

The account of causality outlined here is similar in so many ways to Lewis' account that one may ask where they actually differ. They do differ in many ways but the most important are three. The first was alluded to in subsection 3.1 and concerns the existence of possible worlds. Whereas Lewis believed other worlds exist objectively and independent of thought, I believe they originate in our imagination. Second, Lewis aims to provide a framework that "can serve alike under indeterminism or determinism" (1986d, p. 179). The account outlined here on the other hand has been formulated with indeterminism in mind, and I am not ready to say yet how related they are in the case when the outcome

space only contains a single world, which can be interpreted as a version of determinism. Third, Lewis' account "is in terms of counterfactual conditionals about probability; not in terms of conditional probabilities" (same place, p. 178). Here, conditional probability is one of two aspects of causality and (formal) counterfactual conditionals play no role.

With respect to similarities the most important is how close my ideas regarding causality are to Lewis' (1986a) view on causal explanation—in particular sections I and II. Another similarity concerns the interpretation of probability. Events, that is, elements of \mathcal{F} , are sets of possible worlds, and the conditional probability (say) $P(E|C)$ is the objective propensity of the event C to bring about the event E . In other words, conditional probability may apply to single instances of cases. The propensity (probability) is interpreted in the objective sense as opposed to the subjective, but this should not be interpreted as a critique against subjective accounts of probability. Indeed, in the words of Lewis:

"We subjectivists conceive of probability as the measure of reasonable partial belief. But we need not make war against other conceptions of probability, declaring that where subjective credence leaves off, there nonsense begins. Along with subjective credence we should believe also in objective chance. The practice and analysis of science require both concepts"—(1986e, p. 83).¹⁰

5.4 Conditional expectations re-interpreted

It is common in econometrics to model the impact of one set of variables on another by means of conditional expectations. In its general form such conditional expectations may be denoted $E(X_t|\mathcal{I} = I)$, where X_t is the random variable or variables in question, \mathcal{I} is a σ -field contained in \mathcal{F} and $I \in \mathcal{I}$.¹¹ A common example are so-called "filtrations". If $\mathbf{X}_t = \{X_0, X_1, \dots, X_t\}$ denotes as sequence of the random variable up to and including t , then the sequence of σ -fields generated by $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_t$, commonly denoted $\mathcal{I}_0, \mathcal{I}_1, \dots, \mathcal{I}_t$ and called "information-sets", is the filtration of \mathbf{X}_t if it is the case that $\mathcal{I}_{t-1} \subset \mathcal{I}_t$ and $\mathcal{I}_t \subset \mathcal{F}$ for each t . As a consequence, the conditional expectation $E(X_t|\mathcal{I}_{t-1} = I_{t-1})$ is often referred to as the conditional expectation of X_t on all the information available up to t , and sometimes even *the* true conditional expectation. This is a peculiar practice if the information-set is interpreted to contain what its name suggest, namely information.

The structure of the outcome space permits us to distinguish between two distinct but compatible and complementary ideas, history and information. Let $I_t \in \mathcal{F}$ denote the

¹⁰Essentially this essay is Lewis' account of the relation between subjective and objective versions of probability. See also Lewis (1986c).

¹¹Little is lost by restricting our attention to conditional expectations, since the conditional probability of an event A given and event B is obtained by taking the conditional expectation of the indicator function of A . For instance, $E(I_A|\mathcal{F} = B) = P(A|B)$.

event that an entity (a person, a group of persons, or whatever) possesses or uses a specific piece or pieces of information at t , and let $H_t \in \mathcal{F}$ denote history up to and including t as defined above. Two useful distinctions can be made. Between correct and incorrect information of the past on the one hand, and between complete and incomplete information of the past on the other. More formally, sets of correct and incorrect information are characterised by $I_t \cap H_t \neq \emptyset$ and $I_t \cap H_t = \emptyset$, respectively, and sets of complete and incomplete correct information by $I_t \cap H_t = H_t$ and $I_t \cap H_t \subsetneq H_t$, respectively. This gives three cases. The first case is when the information in the information-set is both correct and complete, and is of course entirely unrealistic. Formally, $I_t = H_t$. The second case is when I_t contains some correct information, but not all the correct information that exists. Formally, $I_t \cap H_t \neq \emptyset$ and $I_t \subsetneq H_t$. Finally, the third case is when I_t contains incorrect information only. Formally, $I_t \neq \emptyset$ and $I_t \cap H_t = \emptyset$.

The point I am driving at is intuitively obvious, namely that in practical econometrics our information is both incomplete and possibly incorrect, and that we use this suboptimal information in estimating conditional expectations. An attempt to formalise this idea could be the following. The "correct" or true expectation conditional on history is given by $E(X_t|\mathcal{F} = H_t)$, whereas what the econometrician in practice estimates is $E(X_t|\mathcal{I}_t = I_t)$ where I_t is an incomplete and possibly partially incorrect information set. Denoting this estimate by $\hat{E}(X_t|\mathcal{I}_t = I_t)$, we may say that one of the key concerns of econometrics is that of efficiently choosing and making use of information such that $\hat{E}(X_t|\mathcal{I}_t = I_t)$ is as close to $E(X_t|\mathcal{F} = H_t)$ as possible.

6 Conclusions

In this essay I have argued that the underlying outcome space in Hendry's (1995) reduction theory can usefully be interpreted as consisting of possible worlds made up of indeterministic and historically inherited particulars. Although the human world is changing all the time in indeterministic ways, the interpretation means reductions can be analysed on the same underlying probability space. This enables several useful concepts and interpretations, of which only a few have been explored in this essay. First, the formulation of theoretical variables can be seen as the "perspective" from which an issue is studied. Second, probabilistic definitions of the absence of measurement error are put forward. Third, a history based probabilistic definition of indeterministic causality that nests discrete, continuous and "interval" versions is proposed. Fourth, mathematical expectation conditional on an information set is re-interpreted. Finally and more generally, a bridge between econometric (/probabilistic) reduction analysis and metaphysics is provided.

This suggests many possible lines for further research, both within the theory and practice of econometrics, only three will be outlined here. First, the unmodified version of Hendry's reduction theory cannot provide an analysis between continuous time theory models and discrete time theory models, whereas the revised version can provide such an analysis, since worlds—by construction—are non-stochastic continuous time processes. For instance, a view that has gained widespread acceptance lately is that volatility forecasts of discrete time models of financial returns should be evaluated against more efficient estimates derived from continuous time theory, see for example Andersen and Bollerslev (1998), Andersen et al. (1999), and Andersen et al. (2005). However, such a practice constitutes a restriction on the discrete time model, since the discrete time models may be compatible with many classes of continuous time models. This leads to the question of how one may analyse the probabilistic reduction such a restriction entails. The notions of subsection 5.4, where conditional expectation on an information-set was re-interpreted, can be used for such a comparison in terms of the distance between $E(\sigma_t|\mathcal{F} = H_t)$, the "true" volatility given by history, and $E(\hat{\sigma}_t^1|\mathcal{I}_t^1 = I_t^1)$ and $E(\hat{\sigma}_t^2|\mathcal{I}_t^2 = I_t^2)$, respectively, one (say) discrete and one (say) continuous time model that condition on different information sets.

Second, the notion of weak stationarity plays a central role in dynamic econometrics and is cast in terms of marginal entities. To recall, a series $\{Y_t\}_{t=1}^T$ is defined as weakly stationary if $E(Y_t) = \mu$ for all t , and given any t we have that $E(Y_t - \mu)(Y_j - \mu) = \sigma_j$ for all j (that is, σ_j does not depend on t). However, everything in the human world is conditional and so it seems natural that this is taken into account in the definition of weak stationarity. Specifically, a straightforward extension is to reformulate the definition of weak stationarity as conditional on a collection of events. Indeed, I would be very surprised if such a definition has not already been proposed and used (although I have been unable to find one in econometric literature). Denote a collection of such conditioning events as $\mathcal{I}^a = \{I_1^a, \dots, I_N^a\}$, where $\mathcal{I}^a \subset \mathcal{F}$ but where \mathcal{I}^a is not necessarily a σ -field. An example of a "conditional" definition of weak stationarity with respect to \mathcal{I}^a would then be that, for all $I^a \in \mathcal{I}^a$, $E(Y_t|\mathcal{F} = I^a) = \mu$ for all t , and given any t we have that $E[(Y_t - \mu)(Y_j - \mu)|\mathcal{F} = I^a] = \sigma_j$ for all j . Breaks or change in stationarity could then be sought explained in terms of a change from one collection of conditioning events \mathcal{I}^a to another set \mathcal{I}^b . For example, a "break" from μ to μ' could be sought explained in terms of a change in circumstances I^a , say, the existence of the Bretton Woods Order of international finance, to I^b , say, the breakdown of the Bretton Woods Order. This would provide a conceptual solution and a unifying framework to study such ideas as co-breaking (Hendry and Massmann 2005), changing "unconditional" volatilities (Engle and Gonzalo

2005; how can changes in unconditional volatility be explained unless the change is due to a change in economic or other circumstances?) and "common features" (Engle and Kozicki 1993).

Finally, the purpose of Hendry's (1995) reduction theory is to provide a probabilistic framework for the analysis of the relation between portions of human reality and empirical econometric models thereof. However, by means of a straightforward extension the relation between abstract economic theories and empirical econometric models can also be analysed probabilistically within the revised reduction theory. The extension consists of augmenting the outcome space Ω with—if needed—the impossible worlds whose subsets all have probability equal to zero, so that the augmented outcome space Ω^+ now contains both the possible worlds and the impossible worlds needed for the study of how theory models and empirical models are related probabilistically. Probabilistic analysis of the relation between abstract theory models and empirical models can then be conducted in terms of the probabilistic definitions of measurement validity in section 4.3, and constitutes a powerful probabilistic alternative to Stigum's (2003) account of the "bridge" between theory and data.

References

- Aigner, D. J. and A. Zellner (1988). Causality. *Supplement to Journal of Econometrics* 39.
- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.
- Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. S. Diebold (2005). Volatility Forecasting. Center for Financial Studies Discussion Paper 2005/08, Frankfurt.
- Andersen, T. G., T. Bollerslev, and S. Lange (1999). Forecasting Financial Market Volatility: Sample Frequency vis-à-vis Forecast Horizon. *Journal of Empirical Finance* 6, 457–477.
- Bauwens, L., H. P. Boswijk, and J.-P. Urbain (forthcoming). Causality and exogeneity in econometrics. *Journal of Econometrics*.
- Cook, S. and D. F. Hendry (1994). The Theory of Reduction in Econometrics. In B. Hamminga and N. B. De Marchi (Eds.), *Idealization IV: Idealization in Economics*, Poznan Studies in the Philosophy of the Sciences and the Humanities 38. Amsterdam: Rodopi B.V.
- Craig, E. (Ed.) (2000). *Concise Routledge Encyclopedia of Philosophy*. London: Routledge.
- Crane, T. (1995). Possible worlds. In T. Honderich (Ed.), *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.

- Crano, W. D. and M. B. Brewer (2002). *Principles and Methods of Social Research*. London: Sage Publications.
- de Vaus, D. (2001). *Research Design in Social Research*. London: Sage Publications.
- Engle, R. F. and J. Gonzalo Rangel (2005). The Spline GARCH Model for Unconditional Volatility and its Global Macroeconomic Causes. Available as <http://pages.stern.nyu.edu/~rengle/SplineGarch.pdf>.
- Engle, R. F., D. F. Hendry, and J.-F. Richard (1983). Exogeneity. *Econometrica* 51, 277–304.
- Engle, R. F. and S. Kozicki (1993). Testing for Common Features. *Journal of Business and Economic Statistics* 11, 369–380.
- Florens, J.-P. and M. Mouchart (1980). Initial and Sequential Reduction of Bayesian Experiments. CORE Discussion Paper 15/1980, Louvain la Neuve (Belgium).
- Florens, J.-P. and M. Mouchart (1985). Conditioning in Dynamic Models. *Journal of Time Series Analysis* 6, 15–34.
- Florens, J.-P., M. Mouchart, and J.-F. Richard (1990). *Elements of Bayesian Statistics*. New York: Marcel Dekker.
- Forbes, G. (1995). Possible Worlds. In J. Kim and E. Sosa (Eds.), *A Companion to Metaphysics*. Oxford: Blackwell Publishers Ltd.
- Geweke, J. (1984). Inference and Causality in Economic Time Series Models. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics: Volume II*. Amsterdam: North Holland.
- Good, I. (1961). A Causal Calculus I. *British Journal for the Philosophy of Science* 11, 305–318.
- Good, I. (1962). A Causal Calculus II. *British Journal for the Philosophy of Science* 12, 43–51.
- Good, I. (1963). Errata and Corrigenda. *British Journal for the Philosophy of Science* 13, 88.
- Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica* 12, iii–vi+1–115. Supplement.
- Hawthorn, G. (1995). *Plausible Worlds*. Cambridge: Cambridge University Press.
- Heil, J. (1998). *Philosophy of Mind. A Contemporary Introduction*. London: Routledge.
- Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.
- Hendry, D. F. and S. Cook (1994). The Theory of Reduction in Econometrics. *Poznan Studies in the Philosophy of the Sciences and the Humanities* 38, 71–100.

- Hendry, D. F. and M. Massmann (2005). Co-Breaking: Recent Advances and a Synopsis of the Literature. Research in Progress, available as <http://www.nuff.ox.ac.uk/Users/Hendry/Papers/cobreakreview05.pdf>.
- Hendry, D. F. and J.-F. Richard (1990). On the Formulation of Empirical Models in Dynamic Econometrics. In C. W. Granger (Ed.), *Modelling Economic Series*. Oxford: Oxford University Press. Earlier published in *Journal of Econometrics* 20 (1982), pp. 3-33.
- Honderich, T. (Ed.) (1995). *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.
- Kim, J. (1996). *Philosophy of Mind*. Boulder CO: Westview Press.
- Kim, J. and E. Sosa (Eds.) (1995). *A Companion to Metaphysics*. Oxford: Blackwell Publishers Ltd.
- Kluge, J. (2004). On the Role of Counterfactuals in Inferring Causal Effects. *Foundations of Science* 9, 65–101.
- Lewis, D. (1986a). Causal Explanation. In *Philosophical Papers. Volume II*. Oxford: Oxford University Press.
- Lewis, D. (1986b). *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- Lewis, D. (1986c). Postscript to 'A Subjectivist's Guide to Objective Chance'. In *Philosophical Papers. Volume II*. Oxford: Oxford University Press.
- Lewis, D. (1986d). Postscript to 'Causation'. In *Philosophical Papers. Volume II*. Oxford: Oxford University Press.
- Lewis, D. (1986e). A Subjectivist's Guide to Objective Chance. In *Philosophical Papers. Volume II*. Oxford: Oxford University Press.
- Loux, M. J. (1998). *Metaphysics. A Contemporary Introduction*. London: Routledge.
- Moravcsik, J. (1995). Potentiality/Actuality. In J. Kim and E. Sosa (Eds.), *A Companion to Metaphysics*. Oxford: Blackwell Publishers Ltd.
- Myrdal, G. (1953). *The political element in the development of economic theory*. London: Routledge.
- Myrdal, G. (1969). *Objectivity in Social Research*. London: Duckworth.
- Parkinson, G. (1995). Philosophy and Logic. In N. Jolley (Ed.), *The Cambridge Companion to Leibniz*. Cambridge: Cambridge University Press.
- Pettit, P. (1993). *The Common Mind. An Essay on the Psychology, Society and Politics*. New York.
- Punch, K. F. (1998). *Introduction to Social Research*. London: Sage Publications.
- Reichenbach, H. (1956). *The Direction of Time*. Berkeley and Los Angeles.

- Ruben, D.-H. (1985). *The Metaphysics of the Social World*. London: Routledge.
- Ruse, M. (1988). *Philosophy of Biology Today*. Albany, NY: State University of New York Press.
- Salmon, W. C. (1993a). Causality: Production and Propagation. In E. Sosa and M. Tooley (Eds.), *Causality*. Oxford: Oxford University Press.
- Salmon, W. C. (1993b). Probabilistic Causality. In E. Sosa and M. Tooley (Eds.), *Causality*. Oxford: Oxford University Press.
- Schumpeter, J. (1949). Science and Ideology. *The American Economic Review* 39, 345–359.
- Searle, J. (1991). *Minds, Brains and Science*. London: Penguin Books.
- Stigum, B. P. (2003). *Econometrics and the Philosophy of Economics: Theory-Data Confrontations in Economics*. Princeton: Princeton University Press.
- Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam.
- Weber, M. (1994). Objectivity and Understanding in Economics. In D. M. Hausman (Ed.), *The Philosophy of Economics*. Cambridge: Cambridge University Press.

Table 1: Starting point, action and the resulting reduction in Hendry's theory associated with the first stage of reduction.

Reduction no.	Starting point and resulting reduction	Action
	The economic mechanism under study: The theory variables $\mathbf{U}^* = (\mathbf{U}_1^*, \dots, \mathbf{U}_T^*)$ defined on the probability space (Ω, \mathcal{F}, P)	Data collection and recording of $\mathbf{U}_t \in \mathbf{U}$, that is, the process of trying to measure the $\mathbf{U}_t^* \in \mathbf{U}^*$ variables
1.	The data generation process (DGP): The data set $\mathbf{U} = (\mathbf{U}_1, \dots, \mathbf{U}_T)$ defined on the transformed probability space $(\Omega', \mathcal{F}', P')$	

Table 2: Starting points, actions and resulting reductions associated with Hendry's theory when the outcome set consists of possible worlds made up of indeterministic and historically inherited particulars.

Reduction no.	Starting points and resulting reductions	Action
	A probability space (Ω, \mathcal{F}, P) , where the outcome-space Ω consists of possible worlds made up of indeterministic and historically inherited particulars	The delineation and definition of a set of theory variables $\mathbf{U}^* = (\mathbf{U}_1^*, \dots, \mathbf{U}_T^*)$
1.	The economic mechanism under study: The theory variables $\mathbf{U}^* = (\mathbf{U}_1^*, \dots, \mathbf{U}_T^*)$ defined on the probability space (Ω, \mathcal{F}, P)	Data collection and recording of $\mathbf{U}_t \in \mathbf{U}$, that is, the process of trying to measure the $\mathbf{U}_t^* \in \mathbf{U}^*$ variables
2.	The data generation process (DGP): A data realisation $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_T)$ of the data variables $\mathbf{U} = (\mathbf{U}_1, \dots, \mathbf{U}_T)$ defined on the probability space (Ω, \mathcal{F}, P)	