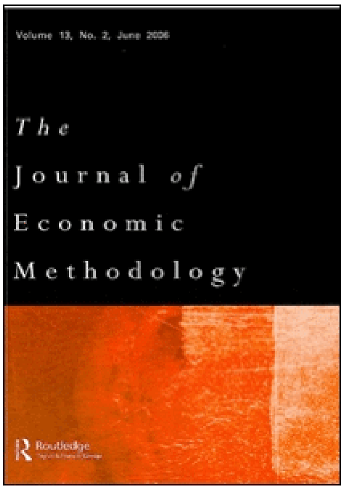

Journal of Economic Methodology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713704064

Econometric reduction theory and philosophy

Genaro Sucarrat [a]
[a] Department of Economics, Universidad Carlos III de Madrid, Spain

Online publication date: 18 March 2010

To cite this Article Sucarrat, Genaro(2010) 'Econometric reduction theory and philosophy', Journal of Economic Methodology, 17: 1, 53 — 75
To link to this Article: DOI: 10.1080/13501780903528978
URL: http://dx.doi.org/10.1080/13501780903528978

Journal of Economic Methodology,
Vol. 17, No. 1, March 2010, 53–75

Econometric reduction theory and philosophy

Genaro Sucarrat*

Department of Economics, Universidad Carlos III de Madrid, Spain

Econometric reduction theory provides a comprehensive probabilistic framework for the analysis and classification of the reductions (simplifications) associated with empirical econometric models. However, the available approaches to econometric reduction theory are unable to satisfactorily accommodate a commonplace theory of social reality, namely that the course of history is indeterministic, that history does not repeat itself and that the future depends on the past. Using concepts from philosophy this paper proposes a solution to these shortcomings, which in addition permits new reductions, interpretations and definitions.

Keywords: theory of reduction; DGP; possible worlds; econometrics and philosophy

JEL Codes: B40; C50

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 objective was to convert more economists to the praxis of evaluating economic theories against economic data using statistical techniques. The deeper question about how the joint n-dimensional probability distribution was related to economic and social reality more generally, however, he remained agnostic about. In his own words, the existence of such a joint probability distribution 'may be purely hypothetical' (ibid., p. iii). Although Haavelmo's ideas had a profound and immediate impact on contemporary economic analysis it nevertheless took until the 1970s and 1980s before the study of the relation between economic reality and models thereof in terms of probability concepts acquired a new momentum. At the centre of several important contributions during these years, including Florens and Mouchart (1980, 1985), Hendry and Richard (1982), Florens, Mouchart, and Richard (1990) and Spanos (1986), was the notion of a 'probabilistic reduction', that is, a probabilistic simplification. In econometrics the term 'reduction' thus has a somewhat different meaning than in philosophy. A probabilistic reduction consists in replacing a complex probabilistic structure with a simpler one – for example through marginalisation and/or conditioning, and a key objective of reduction theory is to study in terms of probability concepts the information that is lost during the simplification process. Reduction analysis has led to the development of important econometric concepts like weak exogeneity, strong exogeneity and super exogeneity, see Engle, Hendry, and Richard (1983), and has been used to justify the widely employed general-to-specific (GETS) methodology for empirical modelling and model evaluation; see Campos, Hendry, and Ericsson (2005) for a comprehensive overview.

*Email: gsucarra@eco.uc3m.es

ISSN 1350-178X print/ISSN 1469-9427 online

DOI: 10.1080/13501780903528978
http://www.informaworld.com

Figure 1 provides a crude overview of the relation between empirical econometric models and social reality, and where probabilistic reduction theory belongs in this relation. It should be underlined that the distinction between three types of econometric models is intended to aid communicating the ideas of this paper. Finer distinctions may of course be more adequate for other purposes, and space limitations do not permit delineating the exact boundaries between the three types. However, examples of each type will be given throughout the paper. The three types of econometric models of social reality outlined in Figure 1 are representation models, estimation and inference models, and empirical models. A simple example of an empirical model is the estimated linear regression:

$$y_t = \hat{a} + \hat{b}x_t + \hat{e}_t \tag{1}$$

where the hat notation suggests the value in question is an estimate. A simple example of its estimation and inference counterpart is the classical regression model:

$$y_t = a + bx_t + e_t, \quad e_t \sim IIN(0, \sigma^2) \tag{2}$$

Assuming the estimation and inference model (2) is a valid representation in some appropriate sense, it can then be used to study to what extent the empirical model (1) is a 'good' depiction of reality in terms of estimator properties, residual properties, proofs, simulations, in-sample and out-of-sample evaluation, and so on. In other words, the study of the relation between empirical models and estimation and inference models corresponds to the traditional conception of theoretical econometrics. Of course, the distinction between empirical models on the one hand and estimation and inference models on the other is not limited to linear, univariate models with a single regressor. Both the empirical

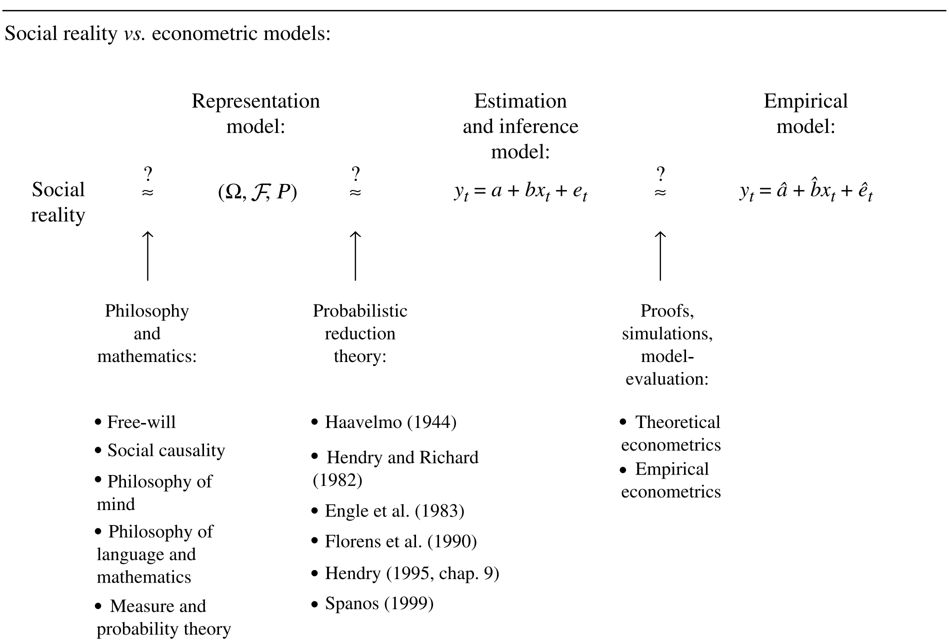

Figure 1. Schematic overview of the relation between social reality and econometric models.

model and the estimation and inference model can be multivariate and/or non-linear in form, in terms of probabilities rather than in terms of conditional expectations, and further distinctions can be made between (say) estimation models on the one hand and inference models on the other, and so on.

Whereas estimation and inference models are intimately related to empirical models, representation models differ in at least two important ways. First, whereas the main purpose of an estimation and inference model is practical in the sense that it is intended to be useful for estimation and/or inference purposes (or alternatively useful for the theory behind estimators and/or inference procedures), the main purpose of a representation model is not practical. Rather, the main purpose of a representation model is to provide a more accurate, detailed and complete depiction of reality from which estimation and inference models can be obtained as probabilistic reductions (simplifications), so that the simplifications and losses of information associated with estimation and inference models can be studied analytically. This leads to the second key difference between estimation and inference models on the one hand and representation models on the other. Whereas estimation and inference models are not necessarily accurate, detailed and complete depictions of reality due to the aim of being useful in econometric practice, representation models constitute relatively accurate, detailed and complete depictions of reality. In Figure 1 a probability space:

$$(\Omega, \mathcal{F}, P) \tag{3}$$

is cited as an example of a representation model, since this in a (narrow) sense is the most general probability structure. In a probability space the outcome set Ω is an arbitrary mathematical set, the event set $\mathcal{F}$ is a σ-field made up of subsets of Ω, and P is a probability measure that designates a value in [0, 1] to each event in $\mathcal{F}$. The probability space constitutes the most general probability structure in a narrow sense, since one instead may start from more general mathematical structures from which, among other things, $\mathcal{F}$ and P can be obtained through reductions.[1] Finally, in addition to measure and probability theory, the study of to what extent probability concepts are capable of accurately depicting social reality draws upon the relevant philosophical literatures, for example the philosophy of language and mathematics literatures, the literature on whether human beings possess a free will, the philosophy of mind literature, and so on.

The ideas and arguments in this paper apply generally to the available approaches to econometric reduction theory. However, in order to make the discussion as specific and relevant as possible for econometric practice in general and economic policymaking in particular, the discussion will be organised in relation to David F. Hendry's (1995, chap. 9) reduction theory. This choice is not restrictive and is based on two observations. First, the starting points of the alternative approaches to reduction theory are either equivalent to or obtainable through reductions of Hendry's starting point the 'economic mechanism'. The economic mechanism, acccording to Hendry, is defined as the 'complete set' of theory 'variables relevant to the economy under investigation' (1995, p. 345). The 'Haavelmo-distribution', the oldest of the approaches, is obtained after five steps of simplification in Hendry's theory (1995, pp. 350–351), whereas Spanos's (1999, p. 3) starting point, the 'stochastic phenomenon', may be interpreted as equal to or a simplified version of Hendry's economic mechanism.[2] The approach of Florens et al. (1990) is Bayesian and mathematically more advanced. However, their treatment is purely technical in the sense that they remain silent about the worldly features that the initial probability structure purports to describe. So analytically their starting point may be seen as equivalent to the underlying probability space in Hendry's theory. The second observation that justifies a focus on Hendry's reduction theory is that Hendry is arguably the most influential contributor to and proponent of the so-called GETS methodology for econometric modelling and model evaluation. The GETS methodology is widely used among economic policymaking and research institutions – see for example Bårdsen, Eitrheim, Jansen, and Nymoen (2005), and Hendry explicitly invokes reduction theory to justify the GETS methodology. The GETS methodology is also known as the 'LSE methodology' after the institution in which it originated, the 'Hendry methodology' after the most influential contributor, and sometimes even 'British econometrics' because the GETS methodology is less popular in the US, see Gilbert (1989), Gilbert (1990), Mizon (1995), Hendry (2003), and Campos et al. (2005).

Although Hendry's reduction theory provides a comprehensive and general framework for the analysis of the relation between social reality and econometric models thereof, it nevertheless has several shortcomings:

(1) Hendry's theory is unable to satisfactorily reconcile two seemingly conflicting views on social reality. The first view is the commonplace theory of social reality that the human world is made up of indeterministic, historically inherited particulars. The exact meaning of this will be explained below in section 3, but crudely it means that the course of history is indeterministic (indeterminism), that history does not repeat itself (particularism), and that the future depends on the past (historical inheritance). The second and seemingly conflicting view is that there are stable laws or regularities regarding the relationship between variables, an idea which underlies most econometric practice. In Hendry's theory the economic mechanism under study, that is, his representation model, is a regularity-entity that *can* change over time. In other words, periods of no change mean the regularities of the economic mechanism are not changing. According to the commonplace theory of social reality, however, there is no *a priori* reason for stable or enduring regularities to exist, so their existence is an empirical question. Conceptually this is not necessarily incompatible with Hendry's theory. But since Hendry does not give a probabilistic account of why and how the economic mechanism changes, his theory is unable to provide probabilistic reduction analysis with reference to the same initial or fundamental probability space. As a solution this paper proposes that the outcome set in the fundamental probability space is specified as consisting of indeterministic worlds made up of historically inherited particulars. This means reduction analysis can be undertaken with reference to the same initial probability space throughout all reductions in Hendry's theory, and the (conditional) existence of regularities and 'true' models – either across time and/or space – can be obtained as (conditional) reductions.

(2) Hendry's theory is in terms of discrete time and can therefore not provide reduction analysis on the relation between continuous and discrete time models. With the proposed structure on the underlying outcome space reduction analysis on the relation between continuous and discrete time models is enabled. Indeed, the relation between events of a wide range of additional temporal structures can be analysed, including intervals, processes, overlapping intervals and processes, and combinations of all of the aforementioned.

(3) According to Hendry there objectively exists a 'complete set' of theory variables 'relevant to the economy under investigation' (Hendry 1995, p. 345). If the course

of history is indeterministic, if history does not repeat itself and if the future depends on the past, the number of theory variables of objective relevance for any economic event is enormous, maybe even infinite. In the words of David Lewis:

> Any particular event that we might wish to explain stands at the end of a long and complicated causal history ... We have the icy road, the bald tire, the drunk driver, the blind corner, the approaching car and more. Together, these cause the crash. Jointly they suffice to make the crash inevitable, or at least highly probable, or at least much more probable than it would otherwise have been ... But these are by no means all the causes of the crash. For one thing, each of these causes in turn has its causes: and those too are causes of the crash. So in turn are their causes, and so, perhaps, *ad infinitum*. (Lewis 1986a, p. 214)

In practice, however, any economic investigation may only focus attention on a (relatively small) finite number of variables that may be of relevance for the purpose of the analysis. Devising the outcome set as consisting of indeterministic worlds made up of historically inherited particulars enables us to treat the formulation or choice of theory variables as a simplification or the perspective from which we study an issue, an idea which in economics has been associated with Weber (1991), Myrdal (1953, 1969) and Schumpeter (1949).

(4) In Hendry's theory the underlying probability space is transformed – again – when data are collected. The theory is therefore unable to provide probabilistic reduction analysis with reference to the same initial probability space of the relation between the theory and data variables. The suggested structure of the fundamental outcome set means the initial probability space does not change and enables a probabilistic definition of the absence of data measurement error.

The proposed structure of the outcome set also enables several new additional reductions, interpretations and definitions, of which only one will be explored: a definition of history is put forward that better conveys the uniqueness and dependence of historical context in probabilistic conditioning on history.

The rest of this paper is organised into five sections. In the next, section 2, the most relevant parts of Hendry's reduction theory for the current purposes are detailed. Section 3 motivates and describes the structure of the outcome space that is proposed. In section 4 the first stage in econometric reduction theory is revisited using the structure of the outcome set proposed in the previous section. Section 5 proposes a definition of history up to time t that more accurately account for historical specificity when conditioning on history, and explores a resulting pair of useful distinctions regarding the relation between history and information. Finally, section 6 concludes.

2 The first stage in Hendry's reduction theory

The purpose of Hendry's reduction theory is 'to explain the origin of empirical models in terms of reduction operations conducted implicitly on the [data generating process (DGP)]' (1995, p. 344), and his theory details 12 reductions associated with various reductive actions whose order is not unique.[3] The 12 reductions are: (1) the data-generation process; (2) data-transformation and aggregation; (3) specification of the parameters of interest; (4) data-partitioning; (5) marginalisation; (6) sequential factorisation; (7) mapping to $I(0)$; (8) conditional factorisation; (9) constancy; (10) lag truncation; (11) functional form approximation; and finally (12) the derived model (Hendry 1995, pp. 360–361 provides a summary). Since the focus in this paper is on the beginning of his theory I concentrate on the first stage 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).[4] 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 paper. Most importantly, random variables and vectors appear in capitals to distinguish them from their realisations, which are denoted in small letters later. 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 $t = 1, \ldots, T$, where the superscript * denotes a perfectly measured variable $\mathbf{U}^* = \mathbf{U}_1^*, \ldots, \mathbf{U}_T^*$, defined on the probability space $(\Omega, \mathcal{F}, P)$... 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 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 $(\Omega, \mathcal{F}, P)$ 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 $(\Omega, \mathcal{F}, P)$, and together $\mathbf{U}^*$ and $(\Omega, \mathcal{F}, P)$ constitute the 'economic mechanism' or 'theory mechanising', that is, the representation model in Hendry's reduction theory. Furthermore, the actions of collecting and recording the data (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), but Hendry's account does not enable much reduction analysis in terms of probability concepts of the relation between $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}, P')$. Schematically the first stage in Hendry's reduction theory is summarised in Table 1.

3 The outcome set as consisting of possible worlds

If $(\Omega, \mathcal{F}, P)$ denotes a probability space where Ω is the outcome set, $\mathcal{F}$ is the event set (a σ-field made up of subsets of Ω) and P is probability measure, then in what follows the

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 $\mathbf{U}^* = (\mathbf{U}_1^*, \ldots \mathbf{U}_T^*)$ defined on the probability space $(\Omega, \mathcal{F}, P)$	
		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 dataset $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_{\hat{T}})$ defined on the transformed probability space $(\Omega', \mathcal{F}', P')$	

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 may be viewed as a probabilistic representation of a social ontology, that is, a probabilistic representation of a theory of the nature of social reality. However, I make no claim to philosophical originality nor to philosophical rigour. The proposed structure is intended to be useful for econometric reduction theory rather than to provide ultimate, irrefutable solutions to philosophical puzzles. For this reason the philosophical discussion and justification is minimal. The proposed structure of the worlds ω is contained in definition 5 in subsection 3.4.

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 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 (1944–2001), 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 l986b, p. 1)

Nevertheless, Leibniz and Lewis differ in several important respects. In particular, whereas Leibniz believed in the objective existence of only a single world, the actual world, Lewis believed in the rather unusual thesis that also 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] This thesis Lewis referred to as 'modal realism'. Lewis's argued in favour of modal realism because, in his view, its advantages outweigh its disadvantages. My own view differs most certainly from Lewis regarding the existence of non-actual worlds, since I only see them as useful mental constructs not existing independent of thought. In this regard, if the reader believes my view gives rise to philosophical issues, it should be recalled that usefulness for econometric reduction theory takes precedence over philosophical rigour, since the ideas proposed here are not intended to provide ultimate, irrefutable solutions to philosophical puzzles.[7]

According to Leibniz and Lewis a world contains everything in the past, present and future. 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 for reduction theory purposes it is not necessary to be all-including backwards and forward in time. What, matters is that the worlds contain everything between a starting point and an ending point, but the portions outside this interval are not really necessary although including them makes little difference. Nevertheless, bounding worlds temporally backwards in time entails an implicit conditioning on the realised history preceding the starting point. Backwards bounding thus means probabilities acquire an interpretation of special interest, but apart from this the only function bounding serves is to simplify the exposition. Henceforth a world ω is therefore devised as a non-stochastic continuous time process:

$$\{s(t) : t \in [0, \infty) \subset \mathbb{R}\}$$

of worldly states-of-affairs $s(t)$ at time t. The initial point $t = 0$ denotes an arbitrary starting point, say, yesterday at midnight or four million years ago, and is not restrictive.[8] For some purposes it is necessary to provide an exact mathematical structure of the states-of-affairs $s(t)$, and one may straightforwardly sketch several such structures. For example, each $s(t)$ may be defined as equal to a countable (finite or infinite) collection of 'attributes', say, $s(t) = \{a_1, a_2, \ldots\}$. In words, a_1 is attribute number 1 of the states-of-affairs, $s(t)$, and so on. This structure is very general and flexible, and accommodates a wide range of ontologies compatible with substance and/or property atomism. A consequence of such a structure is that the most foundational mathematical elements (the 'atoms') of the analysis are properties $\{a_n\}$ that belong to a countable set denoted, say, $\mathbf{a}$.

Interpreting ω as worlds retains the intuitive use of probability algebra. For example, if we would like 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, $A = \{\omega\text{: 10\% unemployed at } t\}$. 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-affairs 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. For example, we may want to devise an event A equal to the set of worlds in which 10% of the labour force of an economy is registered as unemployed over the time interval, say, $[t_0, t_1]$ with $t_0 < t_1$. Or, $A = \{\omega\text{: 10\% unemployed during } [t_0, t_1]\}$.

3.2 Indeterministic 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, in my interpretation, 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 beings are 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 are made up of a finite number of indivisible objects—usually referred to as particles, and on the other hand that human beings are capable of a presumably infinite number of mental states (imagination, thought, and so on).[9] 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 a believer I merely state the thesis as some sort of axiom that I start from. Formally, with respect to the probability space $(\Omega, \mathcal{F}, P)$, indeterminism is simply characterised by Ω containing more than one world ω.

Definition 1. Indeterminism. The worlds $\omega \in \Omega$ are said to be indeterministic if there exists more than one world ω in Ω, and if there exists a pair $\omega \neq \omega'$ such that $\omega \cap \omega' \neq \emptyset$, where $\omega, \omega' \in \Omega$.

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. So Ω must contain more than one world for indeterminism to hold. The intersection property $\omega \cap \omega' \neq \emptyset$ is needed in order to ensure true indeterminism, even when Ω contains more than one world.[10] To see this consider the situation where Ω contains only two worlds. Suppose further that the two worlds do not intersect and that one of the worlds is the actual world. The only possible world is then the actual world, since the other world is not compatible with any part of the actual world. An example of intersecting worlds is worlds that share a common starting point $s(0)$.[11]

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 concerns the states-of-affairs $s(t)$, and intuitively it is the thesis that, literally, history does not repeat itself (no two states-of-affairs are exactly equal in all respects).[12] Formally this may be stated as follows.

> **Definition 2. States-of-affairs particularism**. A world $\omega = \{s(t){:}t \in [0,\infty)\} \in \Omega$ is said to be made up of states-of-affairs particulars if for all pairs $t, t' \in [0,\infty)$, such that $t \neq t'$ and $s(t), s(t') \in \omega$, then $s(t) \neq s(t')$.

In words, two states-of-affairs $s(t)$ and $s(t')$, both of which occur in the same world ω can never be equal in all respect and so $s(t) \neq s(t')$ when $t \neq t'$. However, it should be noted that the definition allows for $s(t)$ and $s(t')$ to be equal in some respects, that is. $s(t) \cap s(t') \neq \emptyset$. For example, if we define states-of-affairs as equal to countable sets of attributes, then $s(t) \cap s(t') \neq \emptyset$ is simply the respects or attributes in which the two states-of-affairs are equal.

3.3 Historically inherited particulars

A further thesis I start from is that the current and the future depend on and inherit characteristics of the past. Differently put, every turn history takes contributes in one or another way to the characteristics of the states-of-affairs of the future. This thesis I shall call 'historical inheritance'. Before providing a formal definition of this property, however, we need the idea of a state-of-affairs process up to (but not including) t.

> **Definition 3. States-of-affairs process up to t**. The process $\omega_t = \{\mathrm{s}(a){:}a < t, t \in (0,\infty)\}$ is said to be a states-of-affairs process up to (but not including) t.

Intuitively, ω_t is simply a history up to (but not including) t. The number 0 is not included in the interval $(0, \infty)$ in order to ensure that a cannot be smaller than 0. This guarantees that ω_t is non-empty and means that at least $s(0)$ is always contained in ω_t. We can now define historical inheritance.

> **Definition 4. Historical inheritance**. The outcome space Ω is said to consist of worlds ω made up of historically inherited particulars if:
>
> (a) All $\omega \in \Omega$ are made up of particulars.
> (b) For all pairs of unequal words $\omega^1, \omega^2 \in \Omega$, that is, $\omega^1 \neq \omega^2$: If $\omega_t^1 \neq \omega_t^2$, then $s^1(t') \neq s^2(t'')$ for all $t', t'' \in (t,\infty)$, where $s^1(t') \in \omega^1$ and $s^2(t'') \in \omega^2$.

In words, if two worlds contains the same history up to t (but not at t), then the states-of-affairs of the two worlds differ from each other in at least one respect at every point in the future.

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 5 below. The definition summarises the ideas so far and provides the starting point for the next sections. The definition may be viewed as a probabilistic representation of a social ontology, but it should be underlined that definition 5 constitutes a *very* general description of the nature of social reality. Indeed, many simpler ontologies – both deterministic and indeterministic – can be obtained as special cases by restricting the worlds ω and the outcome set Ω.

> **Definition 5. Outcome set consisting of indeterministic worlds made up of historically inherited particulars**. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let each $\omega \in \Omega$ be equal to a non-stochastic continuous time process $\{s(t) : t \in [0,\infty)\}$ 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 world ω in Ω and at least two unequal worlds $\omega, \omega' \in \Omega$ intersect: $\omega \cap \omega' \neq \emptyset$ (indeterminism).
> (b) For each $\omega \in \Omega$: For all pairs $t, t' \in [0, \infty)$ such that $t \neq t'$ and $s(t), s(t') \in \omega$, then $s(t) \neq s(t')$ (particularism).
> (c) For each pair of unequal worlds ω^1, $\omega^2 \in \Omega$, that is, $\omega^1 \neq \omega^2$: If $\omega_t^1 \neq \omega_t^2$, then $s^1(t') \neq s^2(t'')$ for all $t', t'' \in (t, \infty)$, where $s^1(t') \in \omega^1$ and $s^2(t'') \in \omega^2$ (historical inheritance).

Crudely, in layman's terms, the first property (a) states that the course of history is indeterministic, the second property (b) states that history does not repeat itself, and the third property (c) states that the future depends on the past. It should be noted though that without conditions on the relation between Ω and $\mathcal{F}$, we may not be guaranteed that $\mathcal{F}$ is a σ-field and that P is a probability function. However, a mathematically rigorous and complete analysis of such conditions is beyond the scope of this paper. Indeed, such an analysis is in itself a form of reduction analysis. Henceforth I simply assume such conditions, if necessary, hold.

3.5 David Lewis compared

The approach to possible worlds outlined above is both similar and different in many ways from David Lewis's ideas, so it may be useful to bring out the main differences and similarities more explicitly. The first main difference was alluded to in subsection 3.1 and concerns the existence of possible worlds. Whereas Lewis held that other (non-actual) worlds exist objectively and independent of thought, a thesis he referred to as 'modal realism', I believe they are fictions in the sense that they exist in our imagination only. Second, Lewis's aim is to provide a framework that 'can serve alike under indeterminism or determinism' (1986d, p. 179). The account outlined here, by contrast, has specifically been formulated with indeterminism in mind, and it is unclear (to me) how related they are in the case when the outcome space Ω only contains a single world, a situation which can be interpreted as a version of determinism. Third, Lewis's account 'is in terms of counterfactual conditionals about probability; not in terms of conditional probabilities' (ibid., p. 178). Here, by contrast, counterfactual conditionals play no formal role, and conditional probability may be interpreted as a measure of the causal 'efficiency' of a conditioning event (the antecedent) to bring about the consequent event, see section 5 below.

With respect to similarities the most important concerns the interpretation of probability. Since events are sets of possible worlds, the conditional probability $P(B|A)$ is interpreted as the propensity of the event A to bring about the event B. In other words, conditional probability applies to single instances of cases, and frequency versions can be obtained through cross-restrictions on the probabilities; see section 4.4 for an idea of how this can be done. Furthermore, I agree with Lewis in interpreting the propensity, that is, the probability, in the objective sense as opposed to the subjective; see Lewis (1986c) for his views on the relation between subjective and objective versions of probability.

4 The first stage revisited

The probability space $(\Omega, \mathcal{F}, P)$ in definition 5 provides the starting point of this section. The purpose of this section is to explain in more detail how and why the shortcomings of the first stage in econometric reduction theory are resolved, and to outline some new reductions.

4.1 The formulation of theory 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, 1969), but similar ideas have been expounded by numerous philosophers and social analysts (for example Max Horkheimer and Jürgen Habermas). Since a world contains everything, the formulation of theory variables $\mathbf{U}^*$ and the associated probability function, denoted P^*, can be treated as reductions that reflect some of these choices.

Consider a set of theory variables $\mathbf{U}^*$ delineated by the investigator.[13] Assuming that $\mathbf{U}^*$ is defined on the probability space $(\Omega, \mathcal{F}, P)$, then the $\mathbf{U}^*$ can be interpreted as the theory variables selected for or considered in an economic investigation. The initial or 'fundamental' $(\Omega, \mathcal{F}, P)$ probability space does not change over time because all change is accounted for by the worlds ω, and the formulation of theory variables can therefore be seen as some sort of reduction or pre-marginalisation with a methodological interpretation. A useful notion for this purpose is the theoretical probability function P^*, which is defined as the probability function associated with the smallest σ-field generated by the events of the theory variables. For example, in the simple case where $\mathbf{U}^*$ is equal to a single theory variable U^* that attains two values u_1^* and u_2^*, then the smallest σ-field is $\mathcal{F}^* = \{\emptyset, A, A^C, \Omega\}$ where $A = \{\omega{:}U^*(\omega) = u_1^*\}$ and $A^C = \{\omega{:}U^*(\omega) = u_2^*\}$.[14] It is always the case that $\mathcal{F}^* \subset \mathcal{F}$, and in this specific example the values of P^* are $P^*(\emptyset) = 0$, $P^*(A) = p_1^*$, $P^*(A^C) = 1 - p_1^*$ and $P^*(\Omega) = 1$. Also, since the probability functions P and P^* can be defined in terms of sets of ordered pairs, we have that $P^* \subset P$. In words, the theoretical probability function P^* provides a probabilistic characterisation of the events $\mathcal{F}^*$ associated with the theory variables, but not of all the possible events in $\mathcal{F}$. Differently put, $\mathcal{F}$ gives a richer characterisation of possibilities than $\mathcal{F}^*$ and the economic mechanism is defined as the theory variables $\mathbf{U}^*$ together with the 'smaller' probability space $(\Omega, \mathcal{F}^*, P^*)$. This probability space is smaller compared with the original probability space $(\Omega, \mathcal{F}, P)$, since $\mathcal{F}^* \subset \mathcal{F}$ and $P^* \subset P$.

The formulation of theory variables can thus be viewed as reflecting the events $\mathcal{F}^*$ that are studied as opposed to the events not studied. This reduction I will refer to as 'pre-marginalisation'. I use the term pre-marginalisation because the term marginalisation has a well-established and well-defined meaning in probability analysis in general and in reduction theory in particular. For example, in Hendry's reduction theory marginalisation leads to reduction number five, see Hendry (1995, chap. 9). The set $\mathcal{F} - \mathcal{F}^*$ can be interpreted as the events or portions of reality that are not studied, and the set $P - P^*$ the associated probabilities. Differently put, $\mathcal{F} - \mathcal{F}^*$ together with $P - P^*$ constitute the information loss associated with the formulation of theory variables. To give an example of how pre-marginalisation is a reduction in the sense that it constitutes the perspective or 'conceptual lenses' we view reality with, consider the delineation of theoretical price and theoretical quantity. In defining these two variables as the object of study, 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. The selection of which portions of reality to study and the way they are depicted in terms of variables can thus be treated as a reduction.

4.2 The DGP

The notion of a DGP is obtained in an analogous manner to the economic mechanism. If we denote $\mathcal{F}^D (\subset \mathcal{F})$ the minimal σ-field $\sigma(\mathbf{U})$ associated with the data variables $\mathbf{U}$, and if we denote the associated minimal probability function for P^D $(\subset P)$, then the DGP is defined as $\mathbf{U}$ together with the probability space $(\Omega, \mathcal{F}^D, P^D)$. Relative to the initial probability space $(\Omega, \mathcal{F}, P)$ one source of the information loss is analogous to that of theory variables, namely $\mathcal{F} - \mathcal{F}^D$ and $P - P^D$. In the case of no data measurement error, this is the only source of information loss. In the more likely case of data measurement error, there may be additional (possibly substantial) sources of information loss. The reason for this is that the data variables $\mathbf{U}$ can have a life of their own and may be entirely unrelated to the theoretical variables they purport to measure. To see this recall that any realisation of the data variables $\mathbf{U}$ 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 the realisation could have been obtained. Similarly, for any series of realisations $\mathbf{u}_1, \ldots, \mathbf{u}_T$ there is a set of possible worlds $\{\omega : \mathbf{U}_1(\omega) = \mathbf{u}_1, \ldots, \mathbf{U}_T(\omega) = \mathbf{u}_T\}$ in which the series of realisations could have been obtained. Whether these sets of worlds correspond to the set of worlds in which the theory-concepts attain is an entirely different question. Their relation can however be readily analysed via the initial probability space by means of suitable concepts. I now turn to this type of analysis.

4.3 Data measurement error

In the methodological literature of the social sciences discussions of data 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 data based 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. Or, differently put, the more satisfactorily the operational definition measures the theoretical definition, the smaller the data measurement error.

The idea of a probability space where the outcome set consists of indeterministic worlds made up of historically inherited particulars enables us to formulate definitions of data measurement error in terms of probabilistic concepts. The purpose of this subsection is to put forward such concepts. To this end, recall that 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 $\mathbf{u}^* = (\mathbf{u}_1^*, \mathbf{u}_2^*, \ldots, \mathbf{u}_t^*, \ldots, \mathbf{u}_T^*)$, with $\mathbf{u}_t^* = (u_{t1}^*, u_{t2}^*, \ldots, u_{ti}^*, \ldots, u_{tI(t)}^*)$ for each t, where the symbolism $I(t)$ means the number of theory variables can vary with t. Furthermore, $u_{t1}^* \in X_{t1}^*, u_{t2}^* \in X_{t2}^*$ and so on for each t, where the $\{X_{ti}^*\}$ are arbitrary sets. Similarly, a realisation of the vector of data variables $\mathbf{U}$ is denoted $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t, \ldots, \mathbf{u}_T)$, with $\mathbf{u}_t = (u_{t1}, u_{t2}, \ldots, u_{tj}, \ldots, u_{tJ(t)})$ for each t, where the symbolism $J(t)$ means the number of data variables can vary with t. Also here $u_{t1} \in X_{t1}$, $u_{t2} \in X_{t2}$ and so on for each t, where the $\{X_{tj}\}$ are arbitrary sets. $J(t)$ can differ from $I(t)$ for any (or all) t. Ideally a definition of measurement validity of $\mathbf{U}^*$ should be sequential and formulated for a sequence of pairs $(\mathbf{U}_1^*, \mathbf{U}_1), (\mathbf{U}_2^*, \mathbf{U}_2), \ldots, (\mathbf{U}_t^*, \mathbf{U}_t), \ldots, (\mathbf{U}_T^*, \mathbf{U}_T)$, 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 $t = 1, 2, \ldots, T$ is straightforward. In what follows I will make use of the probabilistic definition of a measurable variable, which is *not* related to what has been called (data) measurement validity or absence of data measurement error hitherto. This may cause some confusion and the reader is hereby warned. Now, recall the probabilistic definition of a measurable variable:

> **Definition 6. Measurable variable**. Let $(\Omega, \mathcal{F})$ 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 $\{\omega : f(\omega) \in G^*\} \in \mathcal{F}$.

Intuitively Ω^* contains the values of the measurable variable f, and in the case where Ω^* is Euclidean space then f is a random vector. For notational convenience I use the symbolism $f : (\Omega, \mathcal{F}) \rightarrow (\Omega^*, \mathcal{G}^*)$ 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

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

where $\mathbf{X}_\mathrm{t}^* = \mathrm{X}_{\mathrm{t}1}^* \times \mathrm{X}_{\mathrm{t}2}^* \times \cdots \times \mathrm{X}_{\mathrm{tI(t)}}^*$ *and* $\mathbf{X}_\mathrm{t}^* = \mathrm{X}_{\mathrm{t}1} \times \mathrm{X}_{\mathrm{t}2} \times \cdots \times \mathrm{X}_{\mathrm{tJ(t)}}$. The elements of $\mathcal{F}$, $\mathcal{G}_\mathrm{t}^*$ and $\mathcal{G}_\mathrm{t}$ will be referred to as worldly events, theory events at t and data events at t, respectively. Measurement validity of the data event $G_t \in \mathcal{G}_\mathrm{t}$ with respect to the theoretical event $G_t^* \in \mathcal{G}^*$ can now be defined in terms of the extent 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 7. 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 extension from events to variables is more or less straightforward, so for convenience only the definition for measurement validity is provided:

Definition 8. Measurement validity of a data variable. A data variable $\mathbf{U}_t : (\Omega, \mathcal{F}) \rightarrow (\mathbf{X}_t, \mathcal{G}_t)$ is said to be measurement valid if each $G_t \in \mathcal{G}_t$ is measurement valid.

Implicitly the definition makes reference to a theory variable $\mathbf{U}_t^* : (\Omega, \mathcal{F}) \rightarrow (\mathbf{X}_t^*, \mathcal{G}_t^*)$ defined on the probability space $(\Omega, \mathcal{F}, P)$.

In the case where there is no measurement error, the DGP defined by the data variables $\mathbf{U}$ together with the probability space $(\Omega, \mathcal{F}^D, P^D)$ is equal to the economic mechanism, which is given by $(\Omega, \mathcal{F}^*, P^*)$ together with $\mathbf{U}^*$. In this case, there is no information loss associated with the data measurement process. In practice, however, the DGP and the economic mechanism are unlikely to coincide, and the information loss will be a function of the discrepancy between $\mathcal{F}^*$ and $\mathcal{F}^D$. In particular, we can attach probabilities to the events that make up the discrepancy between $\mathcal{F}^*$ and $\mathcal{F}^D$. For example, if the probability of the union of the set that make up the discrepancy is zero, that is, $P[\cup_{n=1}^{\infty}(\mathcal{F}^* - \mathcal{F}^D)] = 0$, then we may say that $\mathbf{U}$ is measurement valid almost surely. Similarly, if $P[\cup_{n=1}^{\infty}(\mathcal{F}^* - \mathcal{F}^D)] \neq 0$ then the probability may be interpreted as the (unconditional) probability of incurring data measurement error.

4.4 The existence of regularities as a reduction

If the course of history is indeterministic, if history does not repeat itself and if the future depends on the past, then there is no *a priori* reason for regularities to exist. Their existence is entirely conditional on spatial and historical specificity. The idea of a probability space where the outcome set consists of indeterministic worlds made up of historically inherited particulars enables us to treat the conditional existence of such regularities, be it in terms of theory or data variables (or both), as a reduction.

To see this consider the simple example of a two period sequence of (say) data variables $\{U_t\}$, $t = 1, 2$, where U_t can attain the two values 1 and 0. If we define the data events as $A = \{U_1 = 1\}$, $A^C = \{U_1 = 0\}$, $B = \{U_2 = 1\}$ and $B^C = \{U_2 = 0\}$, respectively, then the smallest σ-field associated with the data events is $\mathcal{F}^D = \sigma(\{A, A^C, B, B^C\})$. The associated probability function is P^D, that is, $P^D : \mathcal{F}^D \rightarrow [0, 1]$, and the question of interest is to what extent an estimation and inference model P^E represents P^D satisfactorily. In particular, consider the possibility of modelling the sequence $\{U_1, U_2\}$ as an independent and identically distributed (IID) sequence, with p and $1 - p$ denoting the probabilities of 1 and 0, respectively. The joint probabilities of the estimation and inference model would then be given by $P^E(1, 1) = p^2$, $P^E(1,0) = P^E(0,1) = p(1 - p)$ and $P^E(0,0) = (1 - p)^2$, the marginal probabilities are given by $P^E(U_t = 1) = p$ and $P^E(U_t = 0) = 1 - p$, respectively, for $t = 1, 2$, and the conditional probabilities are equal to the marginals due to the IID assumption. The estimation and inference model P^E being a conditionally 'true' representation of P^D can be defined in terms of the implied cross-restrictions of the relation between P^E and conditional P^D. Specifically, denote $\mathscr{C}$ as the family of sets $\{C_1, C_2, \ldots, C_n, \ldots\} \subset \mathcal{F}$ in which the relevant cross-restrictions hold conditional on $C_n \in \mathscr{C}$. Specifically, define $\mathcal{F}^D_{\mathscr{C}}$ as the σ-field generated by the data events together with $\mathscr{C}$, that is, $\mathcal{F}^D_{\mathscr{C}} = \sigma(\{A, A^c, B, B^c\} \cup \mathscr{C})$, and denote $P^D_{\mathscr{C}} : \mathcal{F}^D_{\mathscr{C}} \rightarrow [0, 1]$. Of course, by assumption $\mathcal{F}^D_{\mathscr{C}} \subset \mathcal{F}$ and $P^D_{\mathscr{C}} \subset P$. If we restrict ourselves to events $C_n \in \mathscr{C}$ such that $P(C_n) > 0$,[15] then the most important (in this example) restrictions that would have to be satisfied for P^E to be an almost sure representation of P^D conditional on C_n are $P^D_{\mathscr{C}}(1, 1|C_n) = P^E(1, 1) = p^2$, $P^D_{\mathscr{C}}(1, 0|C_n) = P^D_{\mathscr{C}}(0, 1|C_n) = P^E(1, 0) = P^E(0, 1) = p(1 - p)$ and $P^D_{\mathscr{C}}(0, 0|C_n) = P^E(0, 0) = (1 - p)^2$, and $P^D_{\mathscr{C}}(U_t = 1|C_n) = P^E(U_t = 1) = p$ and $P^D_{\mathscr{C}}(U_t = 0|C_n) = P^E(U_t = 0) = 1 - p$ for $t = 1, 2$. More generally we may say that the regularity P^E exists (almost surely) conditional on each set in $\mathscr{C}$ if all the relevant restrictions hold for each set $C \in \mathscr{C}$. In the special case where $\mathscr{C} = \{\Omega\}$, then $P^D(A) = P^D_{\mathscr{C}}(A|\Omega)$ for each $A \in \mathcal{F}^D_{\mathscr{C}}$.

The probability that the cross-restrictions will hold is $P(\cup_{n=1}^{\infty} C_n)$. The greater (unconditional) $P(\cup_{n=1}^{\infty} C_n)$, the greater (unconditional) generality of the regularity P^E. However, greater unconditional $P(\cup_{n=1}^{\infty} C_n)$ is not necessarily better. Indeed, the key is the appropriateness of each C_n. For example, in many cases it is appropriate to condition on sets of worlds C_n that does not contain, say, the outbreak of World War III or other kinds of events that might reduce the precision. To give an example closer to econometric practice, suppose the error term of a regression is $N(0, 2)$ conditional on C_1 and that a comparable regression's error term is $N(0, 3)$ conditional on C_2 with $C_1 \subset C_2$ and $P(C_1) < P(C_2)$. The generality of the second regression is greater because C_1 is strictly contained in C_2, that is, the second regression holds in more worlds than the second regression. However, the first regression is preferable as long as the worlds of interest for the investigation lay within C_1, since the first regression is more precise in terms of the standard error of the regression.

4.5 Theory models as reductions

A common practice in empirical econometric analysis is to start with a theory model *as if* it were the economic mechanism. With respect to Figure 1, however, a theory model is an estimation and inference model since it is not a sufficiently accurate nor complete enough depiction of social reality to be considered a representation model. An example of starting with a theory model as if it were the economic mechanism is microfoundations, that is, the practice of postulating a disaggregate model, a 'micro' model, then deriving an aggregate model (typically called a 'macro' model) implied by the disaggregate model, before finally estimating the aggregate model subject to the restrictions implied by the disaggregate model.[16] The disaggregate starting model is thus the theory model. Another example is that of evaluating discrete time volatility estimates by comparing them against estimates made up of high-frequency data based on continuous time theory, see among others Andersen and Bollerslev (1998), and Andersen et al. (2003). In this literature a continuous time semi-martingale typically serves as the theory model. Both of these approaches are common in contemporary applied econometrics, the first through so called stochastic dynamic general equilibrium (SDGE) models, the second through the use of realised volatility (and its cousins) as models of volatility. Nevertheless, neither micromodels nor continuous time semi-martingales are equal to the economic mechanism, nor are they sufficiently accurate nor complete enough to be considered as representation models. So one may ask: To what extent do such theory models induce simplifications and other sorts of restrictions? This subsection puts forward some concepts and procedures that shed light on this issue. In brief it is proposed that information loss may occur in at least two ways. First, in assuming that the theory model, denoted P^T, is a 'true' representation, and second in restricting a (conditionally valid) regularity, denoted P^E, to be consistent with the theory model P^T. For example, with respect to the microfoundations approach, P^T would be the micro model whereas P^E would be the macro model. The reductions and associated information losses are summarised in Tables 2 and 3. For the sake of expository simplicity no data measurement error is assumed. In the case of data measurement error, then additional losses of information, simplifications and restrictions would be incurred.

In order to study the reductions that result from assuming that the theory model P^T holds, we may use an approach similar to that of the previous subsection. Let $\{U_t^T\}$ denote the variables of the theory model, and let $\mathcal{C}^T$ be the collection of sets in which P^T is a conditionally valid representation in some appropriate sense. For example, $\mathcal{C}^T$ may be the collection of sets in which P^T is a conditionally valid regularity. The unconditional

Table 2. Summary of the starting points, actions and resulting reductions involved in the revisited first stage of reduction.

Reduction no.	Starting point and resulting reduction	Action	Simplification/ information loss
	A probability space $(\Omega, \mathcal{F}, P)$ where the outcome-space Ω consists of indeterministic worlds made up of historically inherited particulars		
		The delineation and definition of theory variables $\mathbf{U}*$	The events that are excluded from analysis together with their associated probabilities: $\mathcal{F} - \mathcal{F}*$ *and* $P - P*$
1.	The economic mechanism: the theory variables $\mathbf{U}^* = (\mathbf{U}_1^*, \ldots \mathbf{U}_T^*)$ together with $(\Omega, \mathcal{F}*, P*)$, where $\mathcal{F}* \subset \mathcal{F}$ and $P* \subset P$		
		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	Functions of the discrepancy between $\mathcal{F}*$ and $\mathcal{F}^D$, and the discrepancy between $P*$ and P^D. For example, $\mathcal{F}* - \mathcal{F}^D$ *and* $P[\mathrm{U}_{u=1}^{\infty}(\mathcal{F}* - \mathcal{F}^D)]$, or alternatively $\mathcal{F}^D - \mathcal{F}*$ and $P[\mathrm{U}_{u=1}^{\infty}(\mathcal{F}^D - \mathcal{F}^*)]$
2.	The data generation process (DGP): The data variables $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_T)$ together with $(\Omega, \mathcal{F}^{\mathrm{D}}, P^{\mathrm{D}})$, where $\mathcal{F}^{\mathrm{D}} \subset \mathcal{F}$ and $P^{\mathrm{D}} \subset P$		

Table 3. Summary (cont.) of the starting points, actions and resulting reductions involved in the revisited first stage of reduction.

Reduction no.	Starting point and resulting reduction	Action	Simplification/ information loss
		Approximating the conditional probability function of the DGP, given by $P^D_{\mathscr{C}^E}$, by means of an estimation and inference model P^E	The restriction of limiting the analysis to the set $\mathscr{C}^E = \{C_1^E, C_2^E, \ldots\}$ where P^E is approximately equal to $P^D_{\mathscr{C}^E}$, and the discrepancy between $P^D_{\mathscr{C}^E}$ and P^E
3.	The estimation and inference model P^E		
		Restricting the estimation and inference model P^E to be derivable from the theory model P^T	The restriction of limiting the analysis to the set $\mathscr{C}^T = \{C_1^T, C_2^T, \ldots\}$, where P^T is approximately equal to $P^D_{\mathscr{C}^T}$, the discrepancy between $P^D_{\mathscr{C}^T}$, and P^T and the loss associated with $\mathbf{U}_{n=1}^{\infty} C_n^E - \mathbf{U}_{m=1}^{\infty} C_m^T$ and $P(\mathbf{U}_{n=1}^{\infty} C_n^E - \mathbf{U}_{m=1}^{\infty} C_m^T)$
4.	The theory model P^T together with the implied estimation and inference model P^E		

probability of P^T being true is then $P(\cup_{m=1}^{\infty} C_m^T)$, and the implied (unconditional) reductions are thus $\Omega - \cup_{m=1}^{\infty} C_m^T$ and $P(\Omega - \cup_{m=1}^{\infty} C_m^T)$, respectively.

The second way in which information loss or a restriction is induced as a result of assuming that a theory model holds, is the consistency requirement between the two models P^T and P^E, that is, that the latter is derivable from the former, or alternatively that the former is more fundamental. Let $\{\mathbf{U}_t^E\}$ denote the variables of the derivable model, and let $\mathcal{C}^E$ be the collection of sets in which P^E is a conditionally valid representation. The consequence of the consistency requirement is effectively the restriction that both P^T and P^E are assumed to be conditionally valid, and not only one of them. Now, this holds for the intersection between any pair of sets C_m^T, C_n^E, where $C_m^T \in \mathcal{C}^T$ and $C_n^E \in \mathcal{C}^E$. Due to an axiomatic property of probability it is always the case that $P(C_m^T \cap C_n^E) \leq P(C_n^E)$ for any pair C_m^T, C_n^E. In words, the probability of P^E being conditionally valid is always greater or equal to both P^T and P^E being conditionally valid.

Although the assumption of a theory model being true is a probabilistic restriction, this does not always imply that theory models are undesirable. Consider for example the suggestion of Andersen and Bollerslev (1998) that financial volatility estimates of empirical low-frequency discrete time models should be evaluated by comparing them with high-frequency estimates based on continuous time theory. In this case $\{\mathbf{U}_t^T\}$ would be a continuous process with $\{\mathbf{U}_t^E\} \subset \{\mathbf{U}_t^T\}$. In words, $\{\mathbf{U}_t^E\}$ is a sample of $\{\mathbf{U}_t^T\}$. The trade-off thus facing the econometrician is to choose between the restriction, say, $\cup_{n=1}^{\infty} C_n^E \cap \cup_{m=1}^{\infty} C_m^T$ on the one hand that is, the restriction that both the discrete time and continuous time models hold, and $\cup_{n=1}^{\infty} C_n^E$ on the other hand, that is, the probabilistically weaker restriction that only the discrete time model P^E holds. The probability that one of the models holds – regardless of whether the other holds – is always equal to or greater than both being valid, since $P(\cup_{m=1}^{\infty} C_m^T)$ is always greater or equal to $P(\cup_{n=1}^{\infty} C_n^E \cap \cup_{n=1}^{\infty} C_n^T)$. As in the trade-off between the generality of a regularity on the one hand and its precision on the other, in econometric practice the fall in probability as a result of postulating the validity of a theory model would have to be evaluated against any possible gains in efficiency due to the reduction in measurement error; see Bauwens and Sucarrat (2008, subsection 2.1) and Sucarrat (2009) for fuller discussions.

5 History and probabilistic conditionals

If the course of history is indeterministic, if history does not repeat itself and if the future depends on the past, then human decisions take place in a historically unique and dependent context that have a bearing upon decision-making. In time series analysis, however, expectations conditional on a realisation of past variables are incapable of fully conveying this historical specificity. To see this consider the realisation of a discrete time series $I_{t-1} = \{\omega : X_1 = x_1, \ldots, X_{t-1} = x_{t-1}\}$. In words, the $x_1, \ldots, x_{t-1}$ are the realised values of a time series from time 1 to time $t - 1$, and I_{t-1} is the event – the set of worlds – in which this specific realisation can come about. In time series analysis it is common to condition on I_{t-1} when one wants to condition on history up to $t - 1$. However, the set I_{t-1} is too large if one wants to fully reflect the historical specificity of decision-making contexts, since the realisation $x_1, \ldots, x_t$ can come about in counterfactual worlds as well, and not only in worlds that contains actual history. This section proposes a definition of history up to t that better conveys the uniqueness of historical context, and the definition may be viewed as a probabilistic interpretation of Lewis's (1986a, pp. 218–219) '*whole* explanation'. The definition also leads to a useful distinction between two distinct but compatible and complementary types of conditioning events, namely history on the one hand and information on the other.

5.1 History and conditional probability

Let H_{t_1} and E_{t_2} denote a conditioning event and a consequent or explanandum event, respectively, where t_1 either precedes or is contemporaneous with t_2.[17] When defined, that is, when $P(H_{t_1}) > 0$, the corresponding conditional probability $P(E_{t_2}|H_{t_1})$ is thus characterised by two dimensions, 'causal' efficiency and historical possibility. The value between 0 and 1 of the conditional probability refers to the degree of effectiveness of the conditioning event H_{t_1} in bringing about the consequent event E_{t_2} (causal efficiency), whereas historical possibility refers to whether the second event, E_{t_2} is possible at all given the first event H_{t_1}. I will return to these ideas shortly. Now, define history up to t as follows:

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

In words H_t is a set that contains all the worlds that contains the state-of-affairs process ω_t as defined in definition 3, and intuitively H_t is exactly what its name suggests, namely history up to t. When greater than zero, then the probability $P(E_{t_2}|H_{t_1})$ is therefore a measure of the effectiveness of history up to t_1 in bringing about E_{t_2}. The event E_{t_2} at t_2 is said to be historically possible if at least one of its worlds is contained in history H_{t_1}, that is, if $E_{t_2} \cap H_{t_1} \neq \emptyset$. Similarly an event is historically impossible if $E_{t_2} \cap H_{t_1} \neq \emptyset$, since H, by construction contains the set of worlds that contains the course of history up to and including t. It should be noted that we may have $E_{t_2} \cap H_{t_1} \neq \emptyset$ and $P(E_{t_2}| \cap H_{t_1}) = 0$ at the same time, that is, that E_{t_2} is historically possible but probabilistically impossible. Another property of interest is that, when t_1 is a point in time and all the worlds in Ω contain the same starting point $s(0)$, then we have that $P(E_{t_2}|H_{t_2}) \longrightarrow P(E_{t_2})$ when $t_1 \longrightarrow 0$. In words this means the probability of an event E_{t_2} conditional on history up to point t_1 tends to the unconditional probability $P(E_{t_2})$ as t_1 goes to the 'initial' starting time 0 of the worlds ω. The reason for this is that $H_t \longrightarrow \Omega$ as $t \longrightarrow 0$, since – by assumption – all the worlds ω in Ω are possible at the initial starting point $t = 0$.

5.2 History vs. information

Let $\mathcal{I}_{t-1}$ be the σ-field generated by the past variables $\{\mathbf{U}_1, \ldots, \mathbf{U}_{t-1}\}$ where $\mathcal{I}_{t-1} \subset \mathcal{F}$. In dynamic econometrics the conditional expectation $E(\mathbf{U}_t|\mathcal{I}_{t-1} = I_{t-1})$, is sometimes referred to as the conditional expectation of $\mathbf{U}_t$ on *all* the information up to $t - 1$, and sometimes as the conditional expectation of $\mathbf{U}_t$ on *history* up to $t - 1$. To see that $E(\mathbf{U}_t|\mathcal{I}_{t-1} = I_{t-1})$ can neither be conditional on all the information up to t nor on a history that fully conveys the specificity of historical context, let us distinguish between two distinct but compatible and complementary ideas, namely history and information. If I_{t-1} denotes an arbitrary non-empty 'information event', for example a realisation $\{\mathbf{u}_1, \ldots, \mathbf{u}_{t-1}\}$, and if H_{t-1} denotes history as it actually unfolds up to $t - 1$, then two useful distinctions can be made: between correct and incorrect information of the past on the one hand, and between complete and incomplete correct information of the past on the other. More formally, sets of correct and incorrect information arc characterised by $I_{t-1} \cap H_{t-1} \neq \emptyset$ and $I_{t-1} \cap H_{t-1} \neq \emptyset$, respectively, and sets of complete and incomplete correct information by $I_{t-1} = H_{t-1}$ and $I_{t-1} \neq H_{t-1}$, respectively. We may then distinguish between three overlapping cases of interest. The first case is when the information in the information set I_{t-1} is both correct and complete, and is of course unrealistic in econometric practice. Formally, $I_{t-1} = H_{t-1}$. The second case is when I_{t-1} contains some correct information, but not all the (correct) information that exists. Formally, $I_{t-1} \cap H_{t-1} \neq \emptyset$ and $I_{t-1} \cap H_{t-1}$. Finally, the third case of interest is when I_{t-1} contains incorrect information. Formally, $I_{t-1} \cap H_{t-1} \neq \emptyset$ and $I_{t-1} \neq H_{t-1}$.

In econometric practice the information is both incomplete and subject to measurement error (which is not necessarily the same as incorrect information), and this suboptimal information is used in estimating conditional expectations. The 'correct' or true expectation conditional on history up to $t - 1$ is given by $E(\mathbf{U}_t|\mathcal{F} = H_{t-1})$, whereas what the econometrician in practice estimates is $E(\mathbf{U}_{t-1}|\mathcal{I}_{t-1} = I_{t-1})$ where I_t is an incomplete and possibly inaccurate information set. Denoting this estimate by $\hat{E}(\mathbf{U}_{t-1}|I_{t-1} = I_{t-1})$, we may say a key objective of econometrics is that of efficiently choosing and making use of information such that our estimate $\hat{E}(\mathbf{U}_{t-1}|\mathcal{F}_{t-1} = H_{t-1})$ is as close to $E(\mathbf{U}_{t-1}|\mathcal{F} = H_{t-1})$ as possible.

6 Conclusions

This paper has suggested that the initial outcome space in econometric reduction theory can usefully be interpreted as consisting of indeterministic worlds made up of historically inherited particulars. Although the human world is changing all the time in indeterministic ways that have bearing upon the future, the interpretation means that all the subsequent reductions can be analysed relative to the same initial probability space. This resolves some shortcomings in econometric reduction theory and enables several new reductions, concepts and interpretations, of which only a few have been explored here. First, the formulation of theoretical variables can be seen as the perspective from which an issue is studied, an idea which in economics is associated with Max Weber, Joseph Schumpeter and Gunnar Myrdal. Second, probabilistic definitions of data measurement error have been put forward. Third, the existence of regularities have been obtained as a conditionally existent reduction. Fourth, a suggestion of how restrictions implied by theory models can be studied in terms of reductions, including the reductive relation between continuous time and discrete time models, has been put forward. Finally, a definition of history that better conveys the historical specificity and dependence of decision-making contexts when conditioning on the past has been proposed.

At a general level, the ideas put forward in this paper provide a bridge between econometric probabilistic reduction analysis and philosophy. This opens up many possible lines for further research within the philosophy, theory and practice of econometrics. There is already a voluminous philosophical literature that employs the idea of possible worlds to shed light on various philosophical issues, and by providing a bridge between these two literatures econometrics can benefit from these insights — and possibly vice versa.

Acknowledgements

I am indebted to several people for useful comments, suggestions and questions at different stages: the anonymous reviewers, Farooq Akram, Luc Bauwens, Vincent Bodart, Giacomo Bonanno, Geert Dhaene, Juan Dolado, David F. Hendry, Sebastien Laurent, Uskali Mäki, Michel Mouchart, Whitney Newey, Bent Nielsen, Ragnar Nymoen, Fatemeh Shadman, Erik Ø. Sørensen, David Veredas, participants at Forskermøtet 2009, seminar participants at the Depto. de Legica, Historia y Filosofía de la Ciencia at UNED, participants at the INEM 2008 conference, conference participants at $(EC)^2$ 2007, participants at ESEM 2007, seminar participants at the department of economics at Universidad Carlos III de Madrid, seminar participants at the department of economics at the University of Oslo, and participants at the doctoral workshop in economics at Université catolique de Louvain. The research has in part been financed by the European Community's Human Potential Programme under contract HPRN-CT-2002-00232, MICFINMA.

Notes

1. This type of reduction analysis is beyond the scope of this paper.
2. In earlier work Spanos used the term 'actual data-generating process' (actual DGP), that is, the result of the first stage of simplification in Hendry's theory, see Spanos (1986, pp. 19–21 and pp. 661–672), instead of stochastic phenomenon. However, it should be noted that to Spanos the actual DPG/stochastic phenomenon is *not* a probabilistic notion. So, to be precise, I am interpreting a probabilistic representation of the stochastic phenomenon to be equal to or simpler than Hendry's economic mechanism.
3. The 'important point', he says, 'is that empirical relationships must arise from these reductions of the DGP' (ibid., p. 345).
4. Chapter 9 in Hendry (1995) is a revised version of Cook and Hendry (1994), which may be viewed as an evolution of the views in Hendry and Richard (1982). One of the anonymous reviewers raised the concern that the account in *Dynamic Econometric* may not be the most up-to-date *written* account available. My understanding is that it is. Later writing, in my view, in which he seemingly defines the DGP differently, does not treat the first reduction at the same detailed level. My understanding is based on my reading of the literature, on my personal communications with Hendry at conferences, seminars and by email, and on conversations with colleagues who are either responsible for part of the literature or know the literature well.
5. Leibniz, being religious, argued that the world is perfect because among all the possible worlds God must have chosen the most perfect one, a view that was ridiculed by Voltaire in his play *Candide* (Crane 1995). In today's philosophical usage, however, the term usually carries little or no religious connotation.

6. The 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 21 pp. 23–21).
7. For 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).
8. What might be restrictive, though, is representing continuous time by means of real numbers. The issue of which mathematical structure best represents continuous time is, however, beyond the scope of this paper.
9. Entries on 'free will' and 'determinism' are contained in virtually any philosophy or metaphysics dictionary, see for example Honderich (1995) or Kim and Sosa (1995), and usually contain suggestions for further reading. An accessible introduction to the issues is Searle (1991), which is based on the author's BBC lectures. 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).
10. I am grateful to Jesús Zamora for pointing this out to me.
11. An issue raised by one of the reviewers is whether we can include 'impossible' worlds in Ω. For example, this would be of interest in order to conduct reduction analysis of the relation between possible worlds and (impossible) worlds in which economic agents have (say) perfect and complete information. In principle it is possible to augment Ω with impossible worlds. However, it does raise a conceptual issue. In particular, although the impossible worlds would (presumably) be designated a probability of zero, they would nevertheless be possible outcomes, just as any value of a normally distributed variable is possible but a zero-probability outcome. In other words, including impossible worlds leads to a conceptual contradiction: impossible worlds are possible albeit with zero probability.
12. A 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, the formal definition (definition 2) only contains the first interpretation.
13. For expository simplicity the number of theory variables (U^*_{Ti}) is henceforth assumed to be finite for each t. Most of the argument that follows goes through in the case of non-finiteness as well, but non-finiteness gives rise to conceptual and philosophical issues that will not be addressed here.
14. In this example it is assumed that $\{U^* = u^*_1\}$ and $\{\mathbf{U}^* = u^*_2\}$ are mutually exclusive and complete, that is, $\{U^* = u^*_t\} \cap \{U^T = u^*_2\} = \varnothing$ (mutual exclusivity) and $\{U^* = u^*_1\} \cup \{U^* = u^*_2\} = \Omega$ (completeness).
15. This may be justified by the fact that, the sets C_n where $P(C_n) = 0$ are probabilistically unimportant.
16. As pointed out by one of the reviewers, although the macro model is at a more aggregate level than the micro model, aggregation procedures are rarely used in practice.
17. Recall, due to the structure of the underlying outcome space, the subindices t_1 and t_2 can be interpreted as a wide range of temporal structures: Points in time, intervals of time, unions of non-contiguous intervals of time, or combinations of any of these.

References

Andersen, T.G., and Bollerslev, T. (1998), 'Answering the skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts', *International Economic Review*, 39, 885–905.

Andersen, T.G., Bollerslev, T., Diebold, F.S., and Labys, P. (2003), 'Modeling and Forecasting Realized Volatility', *Econometrica*, 72, 579–625.

Bardsen, G., Eitrheim, Ø., Jansen, E.S., and Nymoen, R. (2005), *The Econometrics of Macroeconomic Modelling*, Oxford: Oxford University Press.

Bauwens, L., and Sucarrat, G. (2008), UC3M-Working Paper version: WP 08-18 in the Economic Series. *http://hdl.handle.net/10016/2591*. Forthcoming in the *International Journal of*

Forecasting. 'General to Specific Modelling of Exchange Rate Volatility: A Forecast Evaluation'.
Campos J., Hendry D.F., and Ericsson N.R. (eds.) (2005), *General-to-Specific Modeling* (Vols. 1 and 2), Cheltenham: Edward Elgar Publishing.
Cook, S., and Hendry, D.F. (1994), 'The Theory of Reduction in Econometrics', in *Idealization IV: Idealization in Economics*. Poznan Studies in the Philosophy of the Sciences and the Humanities 38, eds. B. Hamminga and N.B. De Marchi, Amsterdam: Rodopi B.V, pp. 71–100.
Crane, T. (1995), 'Possible worlds', in *The Oxford Companion to Philosophy*, ed. T. Honderich, Oxford: Oxford University Press, p. 707.
Crano, W.D., and Brewer, M.B. (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., Hendry, D.F., and Richard, J.-F. (1983), 'Exogeneity', *Econometrica*, 51, 277–304.
Florens, J.-P., and Mouchart, M. (1980), 'Initial and Sequential Reduction of Bayesian Experiments', CORE Discussion Paper 15/1980, Louvain la Neuve, Belgium.
——— (1985), 'Conditioning in Dynamic Models', *Journal of Time Series Analysis*, 6, 15–34.
Florens, J.-P., Mouchart, M., and Richard, J.-F. (1990), *Elements of Bayesian Statistics*, New York: Marcel Dekker.
Forbes, G. (1995), *A Companion to Metaphysics*, eds. J. Kim and E. Sosa, Oxford: Blackwell 'Possible worlds', pp. 404–405.
Gilbert, C.L. (1989), 'LSE and the British Approach to Time Series Econometrics', *Oxford Economic Papers*, 41, 108–128.
——— (1990), 'Professor Hendry's Econometric Methodology', in *Modelling Economic Series*, ed. C.W. Granger, Oxford: Oxford University Press, pp. 279–303. Previously published in (1986) *Oxford Bulletin of Economics and Statistics*, 8, 283–307.
Haavelmo, T. (1944), 'The Probability Approach in Econometrics', *Econometrica*, 12(Supplement), iii–vi, 1–115.
Hawthorn, G. (1995), *Plausible Worlds*, Cambridge: Cambridge University Press.
Heil, J. (1998), *Philosophy of Mind. A Contemporary Introduction*, London: Routledge.
Hendry, D.F. (1993), *Dynamic Econometrics*, Oxford: Oxford University Press.
——— (2003), 'J. Denis Sargan and the Origins of LSE Econometric Methodology', *Econometric Theory*, 19, 457–480.
Hendry, D.F., and Richard, J.-F (1982), 'On the Formulation of Empirical Models in Dynamic Econometrics', *Journal of Econometrics*, 20, 3–33.
Honderich, T. (ed.) (1995), *The Oxford Companion to Philosophy*, Oxford: Oxford University Press.
Kim, J. (l996), *Philosophy of Mind*, Boulder, CO: Westview Press.
Kim, J., and Sosa, E. (eds.) (1995), *A Companion to Metaphysics*, Oxford: Blackwell.
Kluve, 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* (Vol. II), Oxford: Oxford University Press, pp. 214–240.
——— (1986b), *On the Plurality of Worlds*, Oxford: Basil Blackwell.
——— (1986c), 'Postscript to 'A Subjectivist' Guide to Objective Chance', in *Philosophical Papers* (Vol. II), Oxford: Oxford University Press, pp. 114–132.
——— (1986d), 'Postscript to 'Causation'', in *Philosophical Papers* (Vol. II), Oxford: Oxford University Press, pp. 172–213.
Loux, M.J. (1998), *Metaphysics. A Contemporary Introduction*, London: Routledge.
Mizon, G. (1995), 'Progressive Modeling of Macroeconomic Time Series: The LSE Methodology', in *Macroeconometrics: Developments, Tensions and Prospects*, ed. K.D. Hoover, Kluwer Academic Publishers, pp. 107–170.
Moravcsik, J. (1995), 'Potentiality/Actuality', in *A Companion to Metaphysics*, eds. J. Kim and E. Sosa, Oxford: Blackwell, pp. 405–407.
Myrdal, G. (1953), *The Political Element In The Development Of Economic Theory*. Originally published in Swedish, London: Routledge, 1930.
——— (1969), *Objectivity in Social Research*, London: Duckworth.
Parkinson, G. (1995), 'Philosophy and Logic', in *The Cambridge Companion to Leibniz*, ed. N. Jolley, Cambridge: Cambridge University Press, pp. 199–223.

Pettit, P. (1993), *The Common Mind. An Essay on the Psychology. Society and Politics*, Oxford: Oxford University Press.

Punch, K.F. (1998), *Introduction to Social Research*, London: Sage Publications.

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.

Schumpeter, O. (1949), 'Science and Ideology', *American Economic Review*, 39, 315–359.

Searle, J. (1991), *Minds. Brains and Science*, London: Penguin Books.

Spanos, A. (1986), *Statistical Foundations Of Econometric Modelling*, Cambridge: Cambridge University Press.

——— (1999), *Probability Theory and Statistical Inference*, Cambridge: Cambridge University Press.

Sucarrat, G. (2009), 'Forecast Evaluation of Explanatory Models of Financial Variability', *Economics – The Open-Access, Open-Assessment E-Journal*. *http://www.economics-ejournal.org/economics/journalarticles/2009-8 3*.

Weber, M. (1994), 'Objectivity and Understanding in Economics', in *The Philosophy of Economics*, ed. D.M. Hausman, Cambridge: Cambridge University Press, pp. 69–82.