Earman on Underdetermination and Empirical Indistinguishability

Igor Douven
Department of Philosophy, Utrecht University
E-mail: igor.douven@phil.uu.nl

Leon Horsten
Institute of Philosophy, University of Leuven
E-mail: leon.horsten@hiw.kuleuven.ac.be

October 4, 1998

Abstract

Earman (1993) distinguishes three notions of empirical indistinguishability and offers a rigorous framework to investigate how each of these notions relates to the problem of underdetermination of theory choice. He uses some of the results obtained in this framework to argue for a version of scientific anti-realism. In the present paper we first criticize Earman's arguments for that position. Secondly, we propose and motivate a modification of Earman's framework and establish several results concerning some of the notions of indistinguishability in this modified framework. Finally, we interpret these results in the light of the realism/anti-realism debate.

Keywords: Underdetermination, Empirical Indistinguishability, Realism/Anti-Realism Debate, Confirmation Theory.

1 Introduction

The present paper is intended as a constructive criticism of John Earman's article *Underdetermination, Realism and Reason* (1993). Earman's intention in that article is to argue for a position similar to van Fraassen's anti-realism (35).[1] Towards that end, Earman seeks to establish the following thesis:

T There is a logical asymmetry between observational and theoretical hypotheses, which ensures that scepticism about the latter does not automatically spill over to scepticism about the former (cf. 28–29).

The asymmetry between observational and theoretical hypotheses that Earman wants to exhibit is related to the notion of *empirical indistinguishability.* He distinguishes three notions of empirical indistinguishability, and offers a rigorous framework in which these notions can be investigated. He then formulates several propositions concerning these notions of indistinguishability which, when linked to results from confirmation theory, together support T. He subsequently provides detailed arguments for these propositions in the context of his formal framework.

It will be shown that Earman's arguments for some of the supporting propositions for T are unconvincing. The situation with respect to these claims is more complicated than Earman makes it appear, and a deeper investigation of the notion of empirical indistinguishability is required to determine their truth-value. We will take some steps in the direction of this deeper investigation, and suggest an improvement of Earman's framework. It will be seen that the asymmetry Earman argues for in his paper can be defended in the resulting, modified framework. More importantly, however, our technical results concerning the modified framework will reveal a logical asymmetry between observational and theoretical hypotheses that considerably strengthens the case for Earman's thesis T.

2 Earman's Formal Framework for the Investigation of Empirical Indistinguishability

Earman's framework assumes that all scientific hypotheses are formalized in a *first-order* language.[2] An important distinction is that made between hypotheses in, what Earman calls, the language of the evidence (*observational hypotheses*) and hypotheses that outrun the language of the evidence (*theoretical hypotheses*). The idea is to partition the descriptive vocabulary of the language into two parts, V_O and V_T, and then to take the language of the evidence to consist of all the sentences which, apart from the logical vocabulary, only contain terms belonging to V_O. Earman emphasizes that, although in the context of the realism debate it is natural to think of the subscripts as short for

[1] Unless otherwise indicated, all page references are to Earman (1993).

[2] In our discussion, we will take identity to be a logical symbol that is assigned its usual denotation by all first-order interpretations.

'observational' and 'theoretical', it is inessential where exactly the line is drawn (21).[3]

In Earman's framework empirical (in-)distinguishability is taken to be a property of *classes of hypotheses*. More precisely, it is a property of certain classes of hypotheses that are *mutually exclusive*, *self-consistent*, and *jointly exhaustive*. Let us call a class of hypotheses which meets these conditions a *hypothesis partition*. A hypothesis partition is called *theoretical* if it contains at least one hypothesis that is not logically equivalent to a hypothesis stated exclusively in the observational vocabulary; otherwise it is called an *observational* hypothesis partition.[4] Also, in the examples of hypothesis partitions in Earman's paper the hypotheses are all sentences. This might give the false impression that the framework can only handle finitely axiomatizable theories. This is not so however: there is no difficulty in principle in dealing with (non-finitely) recursively axiomatizable theories. Therefore we propose to distinguish between partitions which contain only finitely axiomatizable hypotheses—called *FA-hypothesis partitions*—and those which contain one or more members which are recursively but not finitely axiomatizable—*NFA-hypothesis partitions*. This distinction will be seen to be relevant to many of our results later on in the paper.

Furthermore, the property of empirical distinguishability accrues to hypothesis partitions only relative to *a class* $\mathcal{M}$ *of models of the language* and an *evidence matrix* $\mathcal{E}$. The class of models can be taken to be the set of all models, or a restricted set thereof (23). If it is taken to be a proper subset of the set of all models, then the requirements of joint exhaustiveness and mutual exclusiveness have to be relativized to this restricted set. An evidence matrix is a class of sentences. No general assumptions are made about evidence matrices, other than that their elements—sometimes referred to as evidence statements—belong to the language of the evidence. One might think that some additional general restrictions on what sentences can go into an evidence matrix are in order or in any case that more should be said about why no further restrictions are imposed. It is, for instance, not evident that *every* sentence of the language of the evidence can be regarded as a possible evidence statement. But more about this later on.

Earman formulates three notions of empirical distinguishability. Since only the second and third of these—abbreviated ED_2 and ED_3—are relevant in both Earman's and our paper, we skip the definition of the first kind of distinguishability.

Definition 2.1: A hypothesis partition $\mathcal{H}$ is ED_2 relative to evidence matrix $\mathcal{E}$ and the class of models $\mathcal{M}$ exactly if for all $H_i, H_j \in \mathcal{H}$, there is a sentence $E_i \in \mathcal{E}$ such that (i) for all $M_i \in \mathcal{M}$: if H_i is true in M_i, then so is E_i, and (ii) there is a $M_j \in \mathcal{M}$ in which H_j is true and E_i false.

[3] Of course this will do little to take away the worries of those who think that we cannot differentiate observational and theoretical hypotheses in a syntactic fashion (cf. for instance van Fraassen 1980:53ff). However, we will for the purposes of this paper grant that such a division can be made. For Earman's defense of the theoretical/observational division, see his (1992), chapter 8, section 2.

[4] Earman does not explicitly mention this last condition, but it is clear that it is intended.

In words: this notion of distinguishability requires the hypotheses to have different consequences in the language of the evidence.

Definition 2.2: A hypothesis partition $\mathcal{H}$ is ED_3 relative to an evidence matrix $\mathcal{E}$ and the class of models $\mathcal{M}$ if and only if for all $H_i, H_j \in \mathcal{H}$ and for all $M_i, M_j \in \mathcal{M}$: if H_i is true in M_i, and H_j is true in M_j, then there is a sentence $E_i \in \mathcal{E}$ such that E_i is true in M_i but not in M_j.

Note that here no mention is made of the hypotheses' empirical *consequences*; for a hypothesis partition to be distinguishable in this sense it suffices when any model in which one of the hypotheses is true differs in some respect expressible by an evidence statement from any model in which one of the other hypotheses is true.

Notions of empirical *in*distinguishability are now straightforwardly defined as the negations of the foregoing definitions (notation: $\mathrm{EI}_i = \neg\mathrm{ED}_i$).[5] Unlike EI_3, EI_2 has a familiar ring to it. Roughly, it says that hypotheses are indistinguishable if they have the same empirical consequences.

One further basic notion has to be introduced, viz. that of a *truth identification method*. A truth identification method for a hypothesis partition $\mathcal{H}$ is a function f from finite sequences of evidence statements σ^n (the superscript indicates the length of the sequence) to $\mathcal{H}$. Intuitively, such a function conjectures on the basis of the evidence obtained so far which of the H_i's $\in \mathcal{H}$ is true. A truth identification method is said to be *reliable for* $\mathcal{H}$ relative to $\mathcal{E}$ and $\mathcal{M}$ if and only if, for all models $M_i \in \mathcal{M}$ and all possible finite evidence sequences σ, there exists an n such that for all $m \geq n : f(\sigma^m) = H_j$ exactly if H_j is true in M_i (see Earman 1993, section 3, for details). A hypothesis partition $\mathcal{H}$ is said to be *underdetermined by the evidence* if and only if there exists, relative to the class of models considered, no reliable truth identification method for $\mathcal{H}$.[6]

3 Earman's Anti-Realist Argument

Earman argues for T on the basis of an example of an observational hypothesis partition (Earman's 'Example 2'; 25) and the following general theorem—to which we shall refer as Theorem E—concerning theoretical hypothesis partitions

[5] Although the notions of (in-)distinguishability are defined for classes of hypotheses, we can by slightly abusing the language but without causing confusion also apply them to pairs of members of such classes, like for instance when we say that those members are EI_2, meaning of course that those two hypotheses have the same consequences in $\mathcal{E}$ so that the partition to which they belong is EI_2. Earman often uses the terms in this way.

[6] Prior to the publication of Earman's paper formal learning theorists had already developed flexible and sophisticated frameworks to deal with questions of truth identification. See among others Kelly and Glymour (1989), Kelly (1992); Kelly (1996) is an extensive survey of the results obtained so far in formal learning theory. Admittedly, these frameworks have a much wider scope than Earman's. In some it is for instance possible to deal with non-axiomatizable theories (cf. Kelly and Schulte 1995), which Earman's cannot handle. Nevertheless, we believe that Earman's relatively simple framework is expressive enough to investigate certain important questions that have been raised in the debate between scientific realists and scientific anti-realists.

(24):

Theorem E: Let there be given an evidence matrix $\mathcal{E}$ that is rich enough to contain for any sentence E_i, singular or quantified, whose descriptive terms belong entirely to V_O (that is, belonging to the language of the evidence), a sentence that is logically equivalent to either E_i or $\neg E_i$. Now suppose that there is an acceptable hypothesis partition $\mathcal{H}$ the members of which are stated in a language that outruns the language of the evidence and which is EI_2 with respect to $\mathcal{E}$. Then $\mathcal{H}$ is EI_3 with respect to $\mathcal{E}$.

In other words, every theoretical hypothesis partition that is EI_2, is also EI_3. Since it follows straightforwardly from the definitions of truth identification method and EI_3 that there cannot be a reliable truth identification method for an EI_3 hypothesis partition, a confirmation-theoretically important consequence of Theorem E is that EI_2 theoretical hypothesis partitions are necessarily underdetermined by the evidence.

'Example 2': Suppose all models in $\mathcal{M}$ have countable domains which are fully named in the sense that for each element in the domain there is a constant $\overline{a}_i$ ($i = 1, 2, 3, \ldots$) that names it. Let $\mathcal{H} = \{\forall x \exists y Rxy, \exists x \forall y \neg Rxy\}$ with $R \in V_O$, and let $\mathcal{E} = \{$finite truth-functional combinations of $R\overline{a}_i\overline{a}_j\}$ with i, j ranging over the natural numbers. Then $\mathcal{H}$ is EI_2—the hypotheses simply have *no* non-trivial consequences in $\mathcal{E}$—but any model (with a domain as specified) of one of the hypotheses must evidently differ from a model of the other in a way that is expressible by some element of $\mathcal{E}$. Hence, $\mathcal{H}$ is ED_3.

Hence Theorem E does not hold for observational hypothesis partitions. In addition, Earman argues that there exists a reliable truth identification method for $\mathcal{H}$.[7] So observational hypothesis partitions are not necessarily underdetermined by the evidence.

Earman's Theorem and Example 2 are jointly meant to show that there is a logical asymmetry between observational and theoretical hypothesis partitions which is of confirmation-theoretic importance. If we accept Earman's arguments for the claim that there are interesting cases of EI_2 theoretical hypothesis partitions (section 9 of Earman's paper; but see section 4.3 below), we here seem to have an argument for a scepticism with respect to the unobservable which does not automatically spill over to a scepticism vis à vis the observable.

However, we do not find Earman's anti-realist argument convincing. Specifically, our complaint is that the validity of Earman's argument for an asymmetry between observational and theoretical hypothesis partitions is dependent on an asymmetry in accompanying evidence matrices. That this is so can be seen as

[7] Earman notes that there is no completely reliable truth identification method for $\mathcal{H}$ in this case but that there does exist an 'almost sure' reliable truth identification method for it. A truth identification method is reliable in this weaker sense if it is reliable provided we can at the outset restrict the models which might model the domain under investigation to some subset of the logically possible ones. As Earman points out, it is, in the case of this particular $\mathcal{H}$, reasonable to settle for this 'almost sure' kind of reliability. This claim may not be uncontroversial, but that is topic for another paper.

follows.

On the one hand, the evidence matrix of Theorem E is *very* large: even multiply quantified sentences with alternating blocks of quantifiers can count as evidence statements. In the proof of Theorem E no use is made of the assumption that the hypothesis partition is theoretical. So one would think that the corresponding version of the theorem for *observational* hypothesis partitions also holds. And so it does. The point is that, *given the evidence matrix that Theorem E assumes*, there simply are no EI_2 observational hypothesis partitions, as a moment's reflection reveals.[8] So the corresponding (observational) version of Theorem E is valid, but vacuous.

On the other hand, the evidence matrix of Example 2 is *very* small: for example, no existential statements occur in it—not even sentences of the form $\exists x Px$ such that $P\overline{a}$ does occur in the evidence matrix for some $\overline{a}$. If we add, for every sentence belonging to the evidence matrix of Example 2, all its existential generalizations, then the hypothesis partition will no longer be EI_2. And it will be argued below that it is *reasonable* to require that these existential generalizations belong to the evidence matrix.

In the face of this, there are two lines that can be taken. Either one argues that in a comparison between theoretical and observational hypothesis partitions it is reasonable to use different evidence matrices, or one argues for some plausible condition(s) all evidence matrices should (minimally) satisfy and investigates the logical differences between both kinds of hypothesis partitions given the so defined class of admissible evidence matrices. We believe that the first strategy will not work: if one wants to make a fair comparison between observable and theoretical hypothesis partitions, then one ought to look for results which hold for a *fixed* class of evidence matrices. So we will opt for the second strategy.

4 Steps towards a Deeper Investigation of Earman's Formal Framework

In accordance with our critical remarks in relation to Earman's Example 2, we first make the constraints to be imposed on evidence matrices somewhat stricter. We then proceed to see what formal results concerning observational and theoretical hypothesis partitions can be obtained with those constraints in place.

4.1 Evidence matrices fixed

An evidence matrix contains the class of all *possible evidence statements*, i.e. the class of all statements that (when true) can be verified on the basis of direct observations. Surely all atomic sentences in the observational vocabulary belong to this set. And in effect Earman requires that all *finite truth-functional combinations* of atomic sentences belong to every evidence matrix.[9] Let us call

[8] There is no way of making the hypotheses mutually exclusive.

[9] Earman does this in a somewhat inelegant way: he allows evidence matrices that are not closed under finite truth-functional combinations, but in the definitions of indistinguishability,

this class *Earman's minimal evidence matrix.* We want to go one step further, and require that every existential generalization of every sentence of Earman's minimal evidence matrix belongs to every evidence matrix. The motivation for this requirement is the same as the one that presumably underlies Earman's closure condition. If $P\overline{a}$ and $P\overline{b}$ are in some possible world supported by direct observations, then $P\overline{a} \wedge P\overline{b}$ is directly supported by the same observations. But if $R\overline{a}\overline{b}$ is supported by a direct observation, then $\exists x R\overline{a}x$ is directly supported by the very same observation. Let us call the wider class that results from this slightly stronger closure condition the *minimal evidence class* (MEC).[10]

Earman wants to take a very *liberal* attitude concerning what goes into evidence matrices (20–21). Therefore he takes the requirement that every evidence matrix contains the minimal evidence matrix as the *only* restriction on admissible evidence matrices: one may, if one likes, take a much larger evidence matrix (as in Theorem E), but one does not have to do so (as illustrated by Example 2). Likewise, we will take it to be the *only* requirement on evidence matrices that they contain the minimal evidence class.

Technically, the minimal evidence class can be described as follows:

Definition 4.1.1: MEC is the smallest set such that:

(i) all atomic sentences in the language of the evidence are in MEC;

(ii) if $\phi, \psi \in$ MEC, then $(\phi \vee \psi) \in$ MEC;

(iii) if $\phi, \psi \in$ MEC, then $(\phi \wedge \psi) \in$ MEC;

(iv) if $\phi\overline{a}_i \in$ MEC (for some $\overline{a}_i \in V_O$), then $\exists x \phi[x/\overline{a}_i] \in$ MEC;

(v) if $\phi \in$ MEC and ϕ is Δ_1, then $\neg\phi \in$ MEC.

It is easily seen that the class of sentences thus defined is the class of observational Σ_1-sentences.[11] The restriction to Δ_1-sentences in clause (v) is of course motivated by our intention to deviate as little as possible from Earman's minimal evidence matrix, and to be (almost) just as liberal in this respect as he is.[12]

We have already mentioned that under the modified closure conditions on evidence matrices we propose, Earman's Example 2 as it stands is no longer acceptable, and when we include the existential statements in the evidence matrix, then the hypothesis partition of that example becomes ED_2. Theorem E still holds. But its scope seems very limited, due to the very peculiar choice of evidence matrix. Ideally, one would want theorems that give information about *a large class of evidence matrices*, or, failing that, about a narrow class of evidence matrices which are argued on philosophical grounds to be reasonable.

it is only their closures under taking finite truth-functional combinations that do real work (21).

[10] Of course this definition of minimal evidence class is always *relative to a language.*

[11] For a definition of the notions Σ_1, Π_1, etc., see any standard textbook on recursion theory, e.g. Rogers (1967).

[12] Omitting the restriction to Δ_1-sentences in clause (v) would collapse MEC into the class of *all* sentences in the language of the evidence.

The results in the remainder of the paper all hold under all evidence matrices containing MEC.

4.2 Observational hypothesis partitions

The main question to be addressed in this subsection is whether it is possible to construct an observational hypothesis partition with the relevant properties of the partition in Example 2, i.e. an $\mathrm{EI}_2/\mathrm{ED}_3$ observational FA-hypothesis partition relative to an evidence matrix containing MEC. For if this can be done, the moral of Earman's Example still stands.

In view of the following proposition, such hypothesis partitions in any case cannot be of the 2-element kind of Example 2:

Proposition 4.2.1: If $\mathcal{H}$ is a 2-element observational FA-hypothesis partition, and MEC $\subseteq \mathcal{E}$, then $\mathcal{H}$ is ED_2 with respect to $\mathcal{E}$.

Proof: Let $\mathcal{H}$ be an arbitrary 2-element observational FA-hypothesis partition. Suppose, towards a reductio, that $\mathcal{H}$ is EI_2. Since $\mathcal{H}$ is an acceptable 2-element partition, $\mathcal{H}$ must be $\{H, \neg H\}$ for some observational hypothesis H. Consider the prenex normal form of H (since $\mathcal{H}$ is an FA-partition this can be done); it will be of the form $\mathrm{Q}_1 \ldots \mathrm{Q}_n P$, with $n \geq 0$, Q_i either of the form $\forall x_i$ or $\exists x_i$ and P quantifier-free. The prenex normal form of $\neg H$ will then be of the form $\mathrm{Q}'_1 \ldots \mathrm{Q}'_n \neg P$, with Q'_i being $\forall x_i$ if Q_i is $\exists x_i$ and vice versa. Now consider $\exists x_1 \ldots \exists x_n P$. Evidently, $H \models \exists x_1 \ldots \exists x_n P$. Since $\exists x_1 \ldots \exists x_n P$ is a Σ_1-sentence in the language of the evidence and hence in $\mathcal{E}$, and since, by assumption, H and $\neg H$ have the same $\mathcal{E}$-consequences, it follows that $\neg H \models \exists x_1 \ldots \exists x_n P$. But this means that $\emptyset \models \exists x_1 \ldots \exists x_n P$, i.e. that $\exists x_1 \ldots \exists x_n P$ is a tautology. By the same reasoning it follows that $\exists x_1 \ldots \exists x_n \neg P$ is a tautology. Now take an arbitrary model M the domain of which has only one element; call this element $\overline{a}$. Then, since both $\exists x_1 \ldots \exists x_n P$ and $\exists x_1 \ldots \exists x_n \neg P$ are tautologies, we must have $M \models P[\overline{a}/x_1, \ldots, \overline{a}/x_n]$ and also $M \models \neg P[\overline{a}/x_1, \ldots, \overline{a}/x_n]$. But there cannot be such a model M. Hence our assumption that $\mathcal{H}$ is EI_2 must be false. But $\mathcal{H}$ was arbitrary, so there can be no EI_2 observational FA-hypothesis partitions with only two members. ❒

All examples of observational hypothesis partitions in Earman's paper are of the 2-element kind our proof deals with. But there is no principled reason not to consider n-element or even infinite observational hypothesis partitions. It might be the case of course that our condition on evidence matrices simply rules out the possibility of EI_2 observational hypothesis partitions in general. This is not so, however, as we shall now show. We will first describe a way of generating finite EI_2 observational FA-hypothesis partitions and then do the same for infinite NFA-hypothesis partitions. Subsequently it will be proved that there are $\mathrm{EI}_2/\mathrm{ED}_3$ observational hypotheses matrices.[13]

[13] The results in the following paragraphs are based on the so-called *theory of partially conservative sentences*, a subbranch of proof theory of arithmetic. For proofs of the basic theorems in this domain, see Bennett (1986), chapter 3. Thanks to Albert Visser for pointing out the relevance of this theory to elementary questions concerning the notion of empirical

First we introduce some terminology. Let $\mathcal{L}_A$ be the first-order language of arithmetic. Let S be Σ_n or Π_n for some n, and let S' be the dual of S (i.e. Π_n if $S = \Sigma_n$, and conversely). And let T be a finitely axiomatizable extension of $\mathrm{I}\Sigma_1$.

Definition 4.2.1: $\mathrm{Cons}(S, T)$ is the set of sentences A such that for all B: if $T + A \vdash B$, and $B \in S$, then $T \vdash B$.

When we let T be $\mathrm{I}\Sigma_1$ and S the set of quantifier-free sentences in the language of arithmetic, then the Gödel-sentence for $\mathrm{I}\Sigma_1$ is a non-trivial example of a sentence belonging to $\mathrm{Cons}(S, T)$ (non-trivial in the sense that it is not entailed by T).

Definition 4.2.2: $\mathrm{NCons}(S, T)$ is the set $\{A | \neg A \in \mathrm{Cons}(S, T)\}$.

Then there is the following:

Theorem 4.2.1 (Solovay):[14] $S \cap \mathrm{Cons}(S', T) \cap \mathrm{NCons}(S, T) \neq \emptyset$.

Consequence 4.2.1: For any n, if the class of Σ_n-sentences (Π_n-sentences) of $\mathcal{L}_A$ is the class of possible evidence sentences, and all the vocabulary of $\mathcal{L}_A$ is observational, then there are finite EI_2 observational FA-hypothesis partitions (relative to this evidence matrix).

Proof: Consider, for simplicity, the case where $n = 1$. If we let $S = \Pi_2$, then the previous theorem says that there must be a $B \in \Pi_2 \cap \mathrm{Cons}(\Sigma_2, T) \cap \mathrm{NCons}(\Pi_2, T)$. Since $\Sigma_1 \subseteq \Sigma_2$ and $\Sigma_1 \subseteq \Pi_2$, it follows that:

- $T + B \vdash A \Rightarrow T \vdash A$ for A any Σ_1-sentence.
- $T + \neg B \vdash A \Rightarrow T \vdash A$ for A any Σ_1-sentence.

This means that $T + B$ and $T + \neg B$ have exactly the same set of evidential consequences. So the hypothesis partition $\{T + B, T + \neg B, \neg T\}$ is an EI_2 observational hypothesis partition. ❐

Consequence 4.2.2: For any n, if the class of Σ_n-sentences (Π_n-sentences) of $\mathcal{L}_A$ is the class of possible evidence sentences, and all the vocabulary of $\mathcal{L}_A$ is observational, then there are infinite EI_2 observational NFA-hypothesis partitions (relative to this evidence matrix).

Proof: Let T be Peano Arithmetic. Take a pair of sentences $B, \neg B$ as in the proof of Consequence 1. Now let $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ be a recursive enumeration of the axioms of T. Let $T[\neg\alpha_n/\alpha_n]$ be the set of first-order consequences of $\{\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \neg\alpha_n, \alpha_{n+1}, \ldots\}$. Then the hypothesis partition

indistinguishability.

[14] Albert Visser pointed out to us that Solovay's theorem goes through even for all extensions of the weak arithmetical theory $\mathrm{I}\Delta_0+\mathsf{EXP}$, which is a FA subtheory of the FA theory $\mathrm{I}\Sigma_1$. Bennett (1986:11) makes the general remark that most of the standard theorems in the theory of partially conservative sentences hold for all extensions of relatively weak fragments of Peano Arithmetic.

$\{T + B, T + \neg B, T[\neg\alpha_1/\alpha_1], \ldots, T[\neg\alpha_n/\alpha_n], \ldots\}$ will be of the required sort. ❒

Theorem 4.2.1 and Consequences 4.2.1 and 4.2.2 are propositions about arithmetical languages. But these languages can also be given a physical interpretation. For instance, they could be interpreted as being about an infinite collection of balls of different sizes, linearly ordered by volume, in an otherwise empty universe.[15] In other words, these results can also be taken to bear on *empirical* theories. It is worth observing that both Consequences are very insensitive to the choice of class of possible evidence statements—presumably Earman would welcome this.

For Earman's purposes the results so far still do not suffice of course: there remains the important question as to whether there exist any EI_2/ED_3 observational hypothesis partitions. We will now show that the answer to this question is (a qualified) *yes*.

Proposition 4.2.2: For any observational hypothesis partition $\mathcal{H}$, if every $M \in \mathcal{M}$ is fully named and MEC $\subseteq \mathcal{E}$, then $\mathcal{H}$ is ED_3 with respect to $\mathcal{M}$ and $\mathcal{E}$.

Proof: Let $\mathcal{H}$ be an observational hypothesis partition. Suppose, for a reductio, it is EI_3. Then there must be $M_i, M_j \in \mathcal{M}$ and $H_k, H_l \in \mathcal{H}$ $(k \neq l)$ such that $M_i \models H_k, M_j \models H_l$ and for all Σ_1-sentences A in the language of the evidence: $M_i \models A \Leftrightarrow M_j \models A$. Since (i) M_i, M_j are fully named and (ii) they satisfy the same Σ_1-sentences, they must have the same diagram (i.e., they satisfy the same set of closed atoms and negations of closed atoms). From this it follows that $M_i \cong M_j$ (since a model's diagram describes it up to isomorphism). But this means that $M_i \models H_l$ and $M_j \models H_k$ and thus that M_i, M_j both satisfy $H_k \wedge H_l$. Since H_k, H_l were assumed to be mutually exclusive, it must be that both models satisfy $\perp$. Hence there can be no EI_3 observational hypothesis partitions. ❒

On the assumption of fully named models we then have the following:

Proposition 4.2.3: There exist EI_2/ED_3 observational hypothesis partitions.

Proof: From Consequence 4.2.1/4.2.2 and the preceding proposition. ❒

Absent the assumption of fully named models, it is quite easy to construct an observational EI_3 hypothesis partition. In fact, the assumption is essential in Earman's Example 2; without it the hypothesis partition in that example is EI_3 (relative to the evidence matrix Earman specifies). On the other hand, it is not true that without the assumption, all observational hypothesis partitions are EI_3. For matrices consisting of non-quantified hypotheses this is evident, but if we require $\mathcal{E}$ to contain MEC, then, e.g., also the partition $\{\exists x Px, \forall x \neg Px\}$ is not EI_3.[16] It is straightforward to prove the general claim that for any $\mathcal{H}$, if

[15] Such interpretations have been explored in the philosophy of mathematics as a way to avoid mathematical platonism. See for instance Field (1980).

[16] The universal quantifier in $\forall x \neg Px$ ranges over all the elements in the domain of a model,

$\mathcal{E}$ is chosen in such a way that all $H \in \mathcal{H}$, except perhaps one, are in $\mathcal{E}$, then $\mathcal{H}$ is ED_3.

The foregoing shows that the logical asymmetry between observational and theoretical hypothesis partitions that Earman claimed to have uncovered can be upheld in the modified framework. However, below it will be seen that there is a more effective strategy available to the anti-realist to argue for thesis T.

4.3 Theoretical hypothesis partitions

Given the definition of theoretical hypothesis partitions in section 2, we obtain:

Proposition 4.3.1:

1. There are 2-element EI_2 theoretical hypothesis partitions.
2. There are 3-element ED_2 theoretical hypothesis partitions.

Proof: Let T be a theoretical predicate, let O be an observational predicate, and let $\overline{a}$ be an observational constant. Then for 1, take $\mathcal{H} = \{\exists xTx, \neg\exists xTx\}$. And for 2, take $\mathcal{H}' = \{\exists xTx, \neg\exists xTx \wedge O\overline{a}, \neg\exists xTx \wedge \neg O\overline{a}\}$. ❐

Comparing Proposition 4.3.1.1 with Proposition 4.2.1 reveals an asymmetry between observational and theoretical hypothesis partitions which might seem to be of confirmation-theoretical relevance. For it shows that the result of Proposition 4.2.1, showing that all 2-element observational FA-hypothesis partitions are ED_2, does not carry over to theoretical hypothesis partitions. The EI_2 theoretical hypothesis partition given here is uninteresting, however, if only because neither of its members has any consequences in the language of the evidence. The question whether there are any interesting EI_2 theoretical hypothesis partitions is currently subject to much controversy.[17] As to $\mathcal{H}'$, it should be noted that the partition is EI_3 so that there exists no reliable truth identification method for it.

Recall that Earman's Theorem E was a conditional claim: *if* there exist any EI_2 theoretical hypothesis partitions, *then* there exist EI_3 theoretical hypothesis partitions. Rather than rely on a controversial premise to argue for the existence of EI_3 theoretical hypothesis partitions, we intend to tackle the question as to whether there exist any such hypothesis partitions straight-away. More precisely, we will show that: (1) all (finite and infinite) theoretical FA-hypothesis partitions are EI_3; (2) all finite theoretical NFA-hypothesis partitions are EI_3.

Proposition 4.3.2: If there are ED_3 theoretical FA-hypothesis partitions with more than two elements, then there are 2-element ED_3 theoretical hypothesis partitions.

---

whether they are named or not. Hence $\exists xPx$ cannot be true in any model that satisfies $\forall x\neg Px$. Since $\exists xPx$ is in MEC, $\{\exists xPx, \forall x\neg Px\}$ is ED_3.

[17] See e.g. Laudan and Leplin (1991), Hoefer and Rosenberg (1994), Leeds (1994), Kukla (1996), Leplin (1997).

Proof: Without loss of generality, let $\mathcal{H}$ be an infinite ED_3 theoretical FA-hypothesis partition with hypotheses $\{H_1, H_2, H_3, \ldots\}$. Then pick an element of $\mathcal{H}$ that is not logically equivalent to an observational hypothesis: let it be H_n. Consider the partition $\mathcal{H}' = \{H_n, \neg H_n\}$. $\mathcal{H}'$ will also be a theoretical hypothesis partition, and it is easy to see that it must be ED_3. ❒

The restriction that $\mathcal{H}$ is an FA-hypothesis partition is essential here, for otherwise we have no guarantee that the element $\neg H_n$ of $\mathcal{H}'$ is recursively axiomatizable. We can drop the restriction if at the same time we require that the partition be finite:

Proposition 4.3.3: If there are n-element ED_3 theoretical NFA-hypothesis partitions, then there are 2-element ED_3 theoretical hypothesis partitions.

Proof: Let $\mathcal{H} = \{H_1, \ldots, H_n\}$ be an ED_3 theoretical NFA-hypothesis partition. Pick again a member of $\mathcal{H}$ that is not logically equivalent to an observational hypothesis: let it be H_1. The set of models that make H_1 false is identical with the set $\mathcal{S}$ of models in which at least (and also at most) one of $H_2, \ldots, H_n$ is true. The theory of $\mathcal{S}$ (i.e., the sentences ϕ such that for all models $M \in \mathcal{S}: \ M \models \phi$), $\text{Th}(\mathcal{S})$ for short, can be seen to be recursively axiomatizable as follows: $\phi \in \text{Th}(\mathcal{S})$ exactly if ϕ is a consequence of each of $H_2, \ldots, H_n$. So if $\phi \in \text{Th}(\mathcal{S})$, then, since $H_2, \ldots, H_n$ are all recursively axiomatizable, ϕ will occur after finitely many steps in the enumerations of $H_2, \ldots, H_n$. As there are only $n-1$ enumerations we have to check, we can for each $\phi \in \text{Th}(\mathcal{S})$ verify in a finite number of steps that it is a theorem of the theory. Hence $\text{Th}(\mathcal{S})$ is recursively axiomatizable. Call one such axiomatization $\neg H_1$. Then $\{H_1, \neg H_1\}$ is an acceptable hypothesis partition. It can easily be seen to be ED_3. ❒

Theorem 4.3.1: If $\mathcal{H}$ is a 2-element ED_3 hypothesis partition, then $\mathcal{H}$ is observational.

Proof: Suppose that $\mathcal{H} = \{H, \neg H\}$ is ED_3. We will show that H (and therefore also $\neg H$) is not theoretical.
We know that, in virtue of (a corollary to) Craig's Theorem, since H, $\neg H$ are recursively axiomatizable, their consequences in the language of the evidence are recursively axiomatizable as well. Let H_O (resp. $(\neg H)_O$) indicate the recursive axiomatization of H's (resp. $\neg H$'s) observational consequences. Then let $\mathcal{H}' = \{H_O, (\neg H)_O\}$. It suffices to show that $\mathcal{H}'$ is a hypothesis partition. For then H is logically equivalent to H_O and $\neg H$ is logically equivalent to $(\neg H)_O$, whereby $\mathcal{H}$ is not a theoretical hypothesis partition. So we have to show that $H_O, (\neg H)_O$ are jointly exhaustive and mutually exclusive.
The former is easy to show. For otherwise $\neg H_O \wedge \neg(\neg H)_O$ is consistent, which is impossible since its negation is implied by H as well as by $\neg H$.
The latter is somewhat harder. Suppose, for a reductio, that $H_O \wedge (\neg H)_O$ is consistent. Then take any maximal consistent extension E of $H_O \wedge (\neg H)_O$ in the observational vocabulary. We will show that E is consistent with H, and also with $\neg H$. Suppose E is inconsistent with H. Then by the compactness of first-order logic, there is an observational sentence $E \vdash E_f$ such that $E_f \vdash \neg H$,

i.e. $H \vdash \neg E_f$. But since the observational consequences of H are axiomatized by H_O, we must have $H_O \vdash \neg E_f$. But this is impossible, for H_O and E_f are together in the maximal consistent set E. In a similar way it can be proved that E is consistent with $\neg H$. But that means that there is a model for H which makes E true, and a model for $\neg H$ which makes E true. These worlds are not separated by an evidence sentence (no matter how the class of possible evidence sentences is chosen), so $\mathcal{H}$ is not ED_3, contradicting the assumption. Conclusion: $H_O \wedge (\neg H)_O$ is inconsistent. ❒

Consequence 4.3.1: If $\mathcal{H}$ is an ED_3 FA-hypothesis partition or a finite ED_3 NFA-hypothesis partition, then $\mathcal{H}$ is not theoretical.

Proof: Immediate from the preceding propositions and theorem. ❒

A question that remains open is whether there are any infinite ED_3 theoretical NFA-hypothesis partitions.

What we have here is clearly stronger than Earman's Theorem E: Consequence 4.3.1 asserts that important classes of theoretical hypothesis partitions are EI_3. In combination with our earlier result that all observational hypothesis partitions are ED_3 (Proposition 4.2.2), this reveals a confirmation-theoretically significant asymmetry between observational and theoretical hypothesis partitions. It provides grounds for scepticism about theoretical hypothesis partitions that does not carry over to observational hypothesis partitions and that allows us to suspend judgment on the question of the existence of genuine scientific examples of EI_2 theoretical hypothesis partitions.

4.4 Relativizing to restricted classes of models

We here briefly discuss two objections that could be raised against drawing philosophical conclusions from results about Earman's framework. First, someone could hold that Earman's restriction of the language in which scientific hypotheses are to be regimented to first-order is not innocuous. It is not at all clear that all of Earman's results continue to hold when we move to a second-order language—which we may after all need to express (up to isomorphism) the structures that we want to talk about (see Shapiro 1991). Second-order logic is not compact. Yet for instance Earman's (1992:151) result that (given the evidence matrix of Theorem E) if a theoretical hypothesis partition is EI_1, then it is also EI_3, essentially uses the compactness of first-order logic. So does our proof of Theorem 4.3.1. Secondly, realists might complain that, apart from empirical 'fit', none of the factors which they regard as relevant to the confirmation of scientific theories (like for instance simplicity, explanatory power, coherence with currently accepted theories) has been taken into account. Such factors may help us choose between theories even if no truth identification method (in the sense defined earlier) can.

Earman's framework has a *partial* defense against both charges built into it. As was noted in section 2, in the application of the notions of ED_1, ED_2, ED_3 one might want to let $\mathcal{M}$ be a *restricted* class of models. It is useful to think of such a restriction as specifying the *background theory* which is taken for granted in the test situation. Viewed from this perspective, considering all possible models means testing in a knowledge vacuum. And the more natural and realistic situation is when we make a host of theoretical presuppositions.

Earman does not require that restrictions on $\mathcal{M}$ should be first-order expressible. Nevertheless, if these restrictions are taken to capture background *theories*, one might want to impose *some* restrictions on what is allowed as a restricted class of models, e.g. that it is expressible by a recursively axiomatizable set of sentences in a sufficiently strong language (the language of second-order logic, for instance). Of course one might still object to the arbitrary requirement that the hypotheses that do not belong to the background theory have to be first-order. We think this is a legitimate complaint but will not go into it here.

Relativization also seems to offer a means for expressing the 'theoretical virtues' in Earman's framework. One could for instance contemplate relativizing to classes of 'structurally simple' models, or—which need not amount to the same thing—to classes of models definable by 'simple' theories. However, to this we may expect anti-realists to reply that: (1) it remains to be seen whether these virtues can be made sufficiently precise to be expressible in terms of relativization; (2) if they can, it is an open question whether the logical asymmetry results supporting a selective scepticism no longer hold; and (3) it is unclear how such realist relativization policies can be justified without invoking some form of dogmatism.

Let us conclude this section with an illustration of how relativization to a restricted class of models can make a difference. For this purpose, consider again Consequence 4.2.2. Peano Arithmetic occurs as an element in the hypothesis partition that is constructed in the proof of that consequence. But if we are a priori certain that the real world is to be found among the models of Peano Arithmetic, then we may consider Peano Arithmetic as a background theory which is beyond scrutiny. Relative to the class of models that is thereby determined, the partition $\{B, \neg B\}$, with B as in the proof of Consequence 4.2.1, is a 2-element EI_2 observational FA-hypothesis partition. So if relativization is allowed, Proposition 4.2.1 does not hold.[18]

5 Conclusion

In this paper we have critically discussed Earman's attempt to defend a version of scientific anti-realism on the basis of results from mathematical logic and confirmation theory. We have argued that Earman's argument fails because it depends at crucial junctures on choices of evidence matrix that seem entirely *ad hoc*.

Our criticism was intended to serve a constructive purpose. We argued for a modification of Earman's framework in which evidence matrices meet certain minimal conditions. With those conditions in place we established various results for observational and theoretical hypothesis partitions. These results

[18] In fact, the proof of Proposition 4.2.1 no longer goes through if we exclude all models with only one element in their domain from $\mathcal{M}$. It would be interesting to know how much relativization is required to obtain actual examples of 2-element EI_2 observational FA-hypothesis partitions.

pointed to a philosophically significant asymmetry between the two kinds of hypothesis partitions: all observational hypothesis partitions are ED_3, whereas important classes of theoretical hypothesis partitions are EI_3. This asymmetry is different from the one Earman was arguing for, i.e. that for theoretical, but not for observational, hypothesis partitions, EI_2 implies EI_3. The significance of Earman's asymmetry for the realism/anti-realism debate depends on what some take to be a questionable assumption, namely the existence of any philosophically interesting EI_2 theoretical hypotheses. In contrast, the asymmetry established in our paper is not dependent on this assumption and thus provides more direct support for Earman's anti-realist thesis T.

Acknowledgements

We are greatly indebted to an anonymous referee for *Erkenntnis* for pointing out various deficiencies in a previous version of this paper and for suggesting numerous improvements. We would also like to thank Herman Roelants, Jean Paul Van Bendegem and Bas van Fraassen for helpful comments and discussions. A special debt is due to Albert Visser, who guided us to and through the logical material that was relevant to the topic of this paper.

References

Bennett, C. (1986) *On Some Orderings of Extensions of Arithmetic.* Doctoral Dissertation, Department of Philosophy, University of Göteborg.

Earman, J. (1992) *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, MIT Press.

Earman, J. (1993) 'Underdetermination, Realism and Reason', *Midwest Studies in Philosophy* XVIII, 19–38.

Field, H. (1980) *Science Without Numbers: A Defense of Nominalism*, Princeton University Press.

Hoefer, C. and A. Rosenberg (1994) 'Empirical Equivalence, Underdetermination, and Systems of the World', *Philosophy of Science* 61, 592–607.

Kelly, K. (1992) 'Learning Theory and Descriptive Set Theory', *Logic and Computation* 3, 27–45.

Kelly, K. (1996) *The Logic of Reliable Inquiry*, Oxford University Press.

Kelly, K. and C. Glymour (1989) 'Convergence to the Truth and Nothing but the Truth', *Philosophy of Science* 56, 185–220.

Kelly, K. and O. Schulte (1995) 'The Computable Testability of Theories with Uncomputable Predictions', *Erkenntnis* 43, 29–66.

Kukla, A. (1996) 'Does Every Theory Have Empirically Equivalent Rivals?', *Erkenntnis* 44, 137–166.

Laudan, L. and J. Leplin (1991) 'Empirical Equivalence and Underdetermination', *Journal of Philosophy* 88, 449–473.

Leeds, S. (1994) 'Constructive Empiricism', *Synthese* 101, 187–221.

Leplin, J. (1997) 'The Underdetermination of Total Theories', *Erkenntnis* 47, 203–215.

Rogers, H. (1967) *Theory of Recursive Functions and Effective Computability*, MIT Press.

Shapiro, S. (1991) *Foundations without Foundationalism. A Case for Second-Order Logic*, Oxford University Press.

van Fraassen, B. (1980) *The Scientific Image*, Clarendon Press.