

This article was downloaded by: [University of Helsinki]

On: 01 August 2013, At: 04:32

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Economic Methodology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rjec20>

Three kinds of 'as-if' claims

Aki Lehtinen ^a

^a Department of Political and Economic Studies/Philosophy ,
University of Helsinki , Finland

Published online: 08 Jul 2013.

To cite this article: Aki Lehtinen (2013) Three kinds of 'as-if' claims, *Journal of Economic Methodology*, 20:2, 184-205, DOI: [10.1080/1350178X.2013.801560](https://doi.org/10.1080/1350178X.2013.801560)

To link to this article: <http://dx.doi.org/10.1080/1350178X.2013.801560>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Three kinds of ‘as-if’ claims

Aki Lehtinen*

Department of Political and Economic Studies/Philosophy, University of Helsinki, Finland

As-if locutions are used (a) in order to indicate that an inaccurate or unrealistic assumption is being made because some inaccuracy or unrealisticness is negligible. This kind of claim has two sub-cases. (a₁) The as-if locution is used to indicate that the as-if claim in itself is inaccurate and that its inaccuracy does not matter for the purposes of the investigation. (a₂) It is used to indicate that claims are made without regard to the causal factors that are assumed to exist but are deemed to be unimportant. As-if locutions may also (b) formulate an accurate behavioural assumption by ascribing intentions or cognitions to an entity in an unrealistic manner or (c) indicate that the modeller is not committed to any particular mental assumptions. The various kinds of claims may be recognised by identifying their underlying ‘attributions’. (a₂), (b) and (c) may be used in formulating an accurate claim.

Keywords: as-if; unrealistic assumptions; attributions; fictions

1. Introduction

Many economists and economic methodologists qualify assumptions with an ‘as-if’ locution. Consider, for example, the following cases.

- (a₁) ‘According to Friedman, business firms behave as if they were fully informed, rational calculators aiming at the achievement of profit maximisation, even though clearly these assumptions do not apply to actual firms’ (Boylan & O’Gorman, 1995, pp. 108–109). ‘Individuals are not optimizers, they are adaptive and only behave “as if” they optimize’ (Kirman, 2006, p. 271).
- (a₂) People behave as if they were only motivated by self-interest.¹
- (b) Voters act as if they have obtained perturbed signals concerning the preference profile and as if they have computed probabilities from these signals using statistical reasoning (Lehtinen, 2007, 2008). ‘People may be induced by social codes of behaviour to act *as if* they have different preferences from what they really have’ (Sen, 1973, p. 258).
- (c) Eve chooses *as though* she were maximising the expected value of a utility function (Binmore, 1998, pp. 360–361).

Why do they use such claims, what do they mean and how do we know whether they are justified?

Using an as-if locution tends to imply that one is making an unrealistic assumption. Such an impression stems in part from the fact that Friedman’s (1953) methodology as a whole has become known as his ‘as-if methodology’, and he has become known as a champion of unrealistic assumptions (see e.g., Røgeberg & Nordberg, 2005). The aim in the paper is to clarify the literature on unrealistic and fictional assumptions. For such a purpose, let us distinguish between *realistic* and *accurate* behavioural claims or

*Email: aki.lehtinen@helsinki.fi

assumptions. An assumption is realistic if it is based on the correct causes of the behaviour, and it is accurate if it correctly describes the behaviour.

I argue that as-if claims are frequently made in order to formulate an accurate assumption (cf. Lehtinen & Kuorikoski, 2007). More specifically, the as-if locution is used in making three different kinds of claim, and only one sub-case of a claim (a_1) may unequivocally be taken to express the idea that an inaccurate assumption is being made. The other cases are better understood as ways of expressing assumptions or making behavioural claims in a pragmatically or semantically correct way. As-if claims in themselves are never realistic in the aforementioned sense because they never *explain* why the entity of interest behaves in the way it does. Nevertheless, I will argue that some as-if claims may help in formulating a perfectly realistic behavioural assumption.

Let us use the term ‘as-if locution’ to denote the linguistic expression ‘as if’. As-if locutions are used in order to (a) indicate that an inaccurate or unrealistic assumption is being made because some inaccuracy or unrealisticness is negligible. This kind of claim has two sub-cases. (a_1) First, the as-if locution is used to indicate that the as-if claim in itself is inaccurate and that its inaccuracy does not matter for the purposes of the investigation. (a_2) Second, it is used to indicate that claims are made without regard to the causal factors that are assumed to exist but are deemed to be unimportant. Here the as-if claim itself is formulated with the locution, and it is claimed to be true enough. As-if locutions also are used (b) to formulate an accurate or a realistic claim concerning the behaviour of something by ascribing intentions or cognitions to it in an unrealistic manner, and (c) to indicate that the modeller is not committed to any particular mental assumptions. I will use the shorthand ‘negligibility-indicating’, ‘false mental ascription’ and ‘non-committal-indicating’ for the first, second and third kind of as-if claims, respectively. Sub-cases under a_1 will be called ‘negligibly inaccurate’ or ‘inaccuracy-indicating’ as-if claims, and sub-cases under a_2 ‘negligibility-indicating accurate’ as-if claims. Each of these uses of the as-if locution is based on different falsity related *reason* why it is used.

Musgrave (1981, p. 385) complained some thirty years ago that the ‘logic of as if statements (what follows from them and what does not follow from them) is terribly unclear’. Some economic methodologists (Lagueux, 1994; Mäki 1992, 1998, 2000, 2003, 2004, 2009) discuss the issue, but their focus is limited either to finding out whether certain kinds of as-if claims are compatible with particular philosophical doctrines, or to interpreting the two protagonists in the as-if discussion, Milton Friedman and Hans Vaihinger.² As a result, the logic of as-if claims in economic practice has not been set out in a comprehensive manner. I endeavour to do that here by providing a way of identifying the different as-if claims and the associated truth claims that they are meant to qualify or specify. My main goal in the present account is thus to sort out the different truth claims that are being made, and to show how they are to be evaluated.

As-if claims and locutions are *never necessary* for describing the assumptions of a model mathematically or for deriving the results. Modellers employ them when they wish to specify what kind of evidence is relevant to the truth status of various assumptions or to specify which claims ought or ought not to be evaluated for truth. Such precision is only called for when the truth status is challenged, or is likely to be challenged. This is also why I discuss claims rather than assumptions: in mathematical models, assumptions can be taken to be propositions expressed in the formal descriptions. As-if claims are not assumptions, they are rather claims *about* assumptions. Strictly speaking, as-if locutions are not necessary even in such methodological discussions because the reasons for using an as-if claim can always be expressed differently.³ Recognising the different kinds of as-if claims helps to

avoid misunderstandings when the locution is used. I propose a simple recognition formula: identify the relevant *attribution* (see Section 2).

The structure of the paper is the following. Section 2 lays out the basic concepts through which the different as-if claims are analysed. Sections 3, 4 and 5 are devoted to cases (a), (b) and (c), respectively. Section 3 has two subsections: Section 3.1 discusses negligibly inaccurate as-if claims, and Section 3.2 negligibility-indicating accurate as-if claims. Section 4 is also divided in two subsections: Section 4.1 presents a case study of (b) concerning signal extraction in an expected utility model of voting and Section 4.2 compares the present account to Uskali Mäki's distinction concerning as-if claims. Section 5 discusses another case: Ken Binmore's account of revealed preferences and representation theorems. Section 6 concludes the paper.

2. As-if claims and attributions

The canonical form of as-if claims is given by the following basic schema (Mäki, 1998):

A behaves as if it were B. (AIC)

An as-if claim is an assertion that has the form (AIC), and which concerns how something behaves; it is a *behavioural claim*.⁴ There is a sense in which the *meaning* of all the different kinds of as-if claim is the same⁵: they specify that A behaves as if it were B. Note that (AIC) does not make any claims about whether A is B. It merely states that A behaves as B would behave were A, in fact, B. Let us call the proposition that 'A is B' the underlying *attribution* of an as-if claim, and the proposition that 'A behaves as if it were B' the *as-if claim* (AIC).⁶ The attribution is not a claim at all because the modeller never intends to assert that it is true. Yet, a mathematical model including an assumption that is justified with an as-if locution represents A as behaving like B. The point of using the locution is to focus attention on the fact that A is not asserted to be B, even though the model describes A as behaving like B. The reason why an assumption is qualified with an as-if locution is that it provides understanding about the relevant truth claims of the model. Something is also claimed to be correct in the model. In the case of inaccuracy-indicating as-if claims, that something is the model result, and in all other cases it is the behavioural assumption itself (i.e. the as-if claim).

As Hindriks (2006) notes, presenting assumptions in a meaningful way *presupposes* certain things about the relationship between the model and the world. In (AIC) the presumption is that A is not in fact B. If A were B, there would not be any reason to use the as-if locution because one would simply be able to say 'A behaves like B because it is B'. The idea of a presupposition requires that there must be a *reason* for using the as-if locution. *An agent who uses it must presuppose that the corresponding attribution is false or involves a category mistake. Alternatively, if the locution is used when the attribution could be true, the agent wishes to indicate that he or she is not committed to its truth. Otherwise its use is not warranted.* What is common to all attributions is thus that the agent is not committed to their truth. Finding the attribution for any given as-if claim allows one to determine whether using the as-if locution is appropriate. If it is, one could say that the pragmatic *constraints* for using it are satisfied.

When the as-if claim is accurate it is not usually so in absolute terms, but may be approximately correct: it may be 'true enough'. The intended meaning is that although actual behaviour is not quite instantiated by the features of behaviour in the model, it is not very far from it. Where the divergence is negligible the as-if claims, although not strictly true of the phenomena they denote, are true enough (Elgin, 2004, 2009). Furthermore,

given that satisfying the constraints implies that A is not B, A cannot behave as B in all respects, and may only do so in some contextually salient respect. As-if claims are thus always made relative to the purpose of the investigation. The divergence from the truth of a behavioural claim (AIC) may be appropriate for one purpose but unacceptable for another. It is this pragmatic nature of as-if claims that justifies the need for a practice-based rather than a purely semantic account.

I am now able to state my point more precisely: the as-if claim is inaccurate or false under sub-case (a_1), and may be accurate or even realistic under the other cases. Whether it is indeed true depends not on whether A is B but on whether A's behaviour corresponds to that which it would exhibit if it behaved like B. This provides a *criterion* for evaluating the truth of as-if claims. My aim in the later sections is to demonstrate that there are cases in which A *in fact behaves* as if it were B in the relevant respects even though it clearly is not B.

Identifying different kinds of attributions is also a way of distinguishing between the different kinds of as-if claims on the basis of the primary reason for using the locution. Identifying the right kind of attribution thus facilitates understanding of what kind of truth claim is being made. This is the main benefit to be derived from a classification of as-if claims.

There are two dimensions along which the function and truth status of assumptions may be evaluated that are systematically related to the different kinds of as-if claims. However, the three-part classification is not based on them. First, a major reason for distinguishing between attributions and as-if claims is that the latter can be classified on the basis of whether evaluating the truth of the former provides relevant information concerning their truth. This does not provide a truth criterion, but rather indicates what *kind of evidence* is relevant for evaluating the truth. Secondly, adding an as-if locution to a sentence describing a behavioural claim may or may not change the extent to which the meaning of the attribution and the as-if claim is the same. If they do not have the same meaning, I would say that introducing the locution *brings about a semantic change*.

3. Negligibility-indicating as-if claims

3.1. Negligibly inaccurate as-if claims

The first use of the as-if locution involves making inaccurate assumptions. The primary reason for using it here is that the modeller knows that he or she is making an inaccurate assumption. The simplest way of conveying the idea is to say that A is described as being or behaving like B in the model even though it is not so in the real world. One then says that A behaves as if it were B in the model. This category of as-if claims is closely related to Vaihinger's notion of a fiction; there is a deviation from reality (which may ultimately be eliminated) and an awareness that the fiction is just a fiction.

Let us now have a closer look at Kirman's example of such as-if claims: the as-if claim can be written briefly as follows.

(a_1) People behave as if they optimize.

Kirman makes it clear that although people do not behave optimally, the assumption is made anyway.⁷ The reason why the as-if locution is used is that the modeller knows that the inherent assumption is false. The attribution is 'actual people optimize'. Optimising refers to the endeavour of pursuing optimal results here. The endeavour may not be sufficient for optimal behaviour because the individuals may not have the requisite information, knowledge, foresight etc. in order to succeed. Yet, even if they had the

requisite information but did not try to optimize, one would have to use the as-if locution because the attribution is false irrespective of whether they have the requisite information.

Consider now whether the use of the as-if locution brings about a semantic change in (a_1). This requires specification of what exactly would be claimed if it were not used. An obvious response is that without the locution the claim is about the truth of the attribution. There does not seem to be a semantic change under (a_1) because ‘actual people optimize’ makes the same claim *about behaviour* as the as-if claim ‘people behave as if they optimize’. Although the locution is introduced in order to indicate that the attribution and the behavioural claims are false, the meaning of the attribution and the as-if claim is the same.⁸

One might object that it does not make any sense to claim, for example, that real people are rational if one does not believe that they are. Nevertheless, certain scholars think that there is no problem. Consider, for example, Boland (1979, pp. 512–513):

[Friedman’s,] as if theory of explanation... [I]f we are trying to explain the *effect* of the assumed behavior of some individuals (e.g., the demand curve derived with the assumption of maximizing behavior), *so long as the effect is in fact observed and it would be the effect if they were in fact to behave as we assume*, we can use our behavioral assumption even when the assumption is false. That is, we can continue to claim the observed effect of the individuals’ (unknown but assumed) behavior is *as if* they behaved as we assume. Note carefully, the individuals’ *behavior* is not claimed to be *as if* they behaved as we assume, but rather it is the *effect* of their behavior that is claimed to be *as if* they behaved according to our assumption. (all emphases in the original)

One way of making sense of this is by noting that modellers are not committed to the truth of all assumptions to the same degree. In particular, a modeller who asserts (a_1) is not committed to the truth of the as-if claim itself. If the model result is demonstrably robust with respect to individual behaviour (Kuorikoski, Lehtinen, & Marchionni, 2010; Lehtinen & Kuorikoski, 2007), such behaviour need not provide the central causal mechanism in a rational-choice model (see also Satz & Ferejohn, 1994). One might thus wish to assert (a_1) just because one’s model needs some connection between the individuals and the social-level outcomes. Asserting (a_1) could then be rephrased as ‘the optimising model gives me a connection between individuals and outcomes. Even though real people are not like the individuals in my model, it does not matter because the fine details of their behaviour are irrelevant anyway’. One could thus assert (a) even though one knows that real individuals do not optimize. Another possibility is that although individual behaviour is not irrelevant to the results, it is sufficient if part of the population optimizes.

Note that if such justifications are appropriate, the modeller may well not be committed to any particular determinants of behaviour or be indifferent towards them and their truth. This may make it difficult to distinguish between inaccuracy-indicating (a_1) and non-committal-indicating claims (c). Here the modeller believes that the as-if claim *itself* is false, but for the purposes of his or her model, it is negligibly so. The truth of the claim is not important to the modeller because he or she is merely using the behaviour specified *in order to study something else*, and the as-if locution is used mainly to indicate that one should not evaluate the model by examining the particular assumption that it has qualified. Thus the modeller may believe that optimisation entails behaviour that is different from the way in which people actually behave, but the way in which they behave is ultimately irrelevant to the results of the model. In contrast, in case (c) the modeller aims to formulate an accurate rather than a negligibly false behavioural assumption.

3.2. Negligibility-indicating accurate as-if claims

Here A is some system of interest, and B refers to some set of ideal conditions. Mäki (1992, 2003, 2009) provides the following description of the function of such usage:⁹

(a₂) Phenomena behave as if certain ideal conditions were met: conditions under which only those real forces that are theoretically isolated are active.

Here the as-if claim refers to the idea that various disturbing factors found in the real world are assumed to be absent. The central forces identified in the model are depicted as if they were the only forces. The attribution is that reality satisfies the ideal conditions, i.e. that there are no disturbing factors other than those identified in the model. Introducing the as-if locution induces a semantic change because the sentence now makes a true enough claim about the behaviour of something in actual conditions rather than a false claim concerning whether the reality exhibits ideal conditions. The constraints are satisfied because we know that the ideal conditions do not hold in the real world. Furthermore, the as-if claim directs attention to what is achieved in assuming the absence of such factors: we get to know what kind of behaviour or outcomes would result if the central forces were acting alone. ‘People behave as if they were only motivated by self-interest’ provides a well-known example from economics. In this example the negligible causal factors are mental states, but as we will now see, they also can be physical forces.

The as-if claim expressed by (a₂) may be accurate or true enough. Consider, for example, the often used example of Galileo’s law of falling bodies, which states that ‘under vacuum, the distance covered by a falling object is given by $s = \frac{1}{2}gt^2$ ’. One could then say, with Friedman, that good predictions may be obtained by assuming that

(a₂′) Actual objects fall as if they were following Galileo’s law.

The as-if claim is true enough if the air pressure is sufficiently close to zero, and the object satisfies certain characteristics such as having sufficient density. The attribution is that the air pressure is zero. Given that it must be false, one does not need to use the as-if locution at all if one has artificially created a vacuum, and makes statements concerning objects falling in it. There is, however, a range of conditions under which it makes sense to say that ‘an object behaves as if those conditions corresponded to a vacuum’.

Here, evaluating the truth of the attribution is relevant for evaluating the truth of the as-if claim, which is a feature that (a₁) and (a₂) share. This does not seem to be a case in which one would say ‘A behaves as if it were B even though it *clearly* is not B’ because whether A (an object in actual atmospheric conditions) behaves in such a way depends on how close A is to being B (an object in a perfect vacuum). As I will demonstrate in later sections, however, other kinds of as-if claims are different in this respect.

Friedman (1953, p. 18) emphasises that it is *not assumed* that we live in a vacuum.¹⁰ We merely assume (or claim) that objects fall as if they were in a vacuum. Introducing the as-if locution brings about a semantic change that transforms the false claim about there being a vacuum into a true enough claim about behaviour. Musgrave (1981) suggests that the as-if locution could be used to express a negligibility assumption: the claim ‘air resistance is negligible’ could be expressed in the words ‘bodies fall as if they were in a vacuum’ (see also Mäki, 2000). These two claims are not identical, however, because the former expresses reasons why the behavioural claim might be accurate whereas the latter provides a descriptive claim about the behaviour of something.

Consider now another well-known example, Boyle's law of ideal gases states, roughly, that at a constant temperature the pressure of a fixed amount of gas varies inversely with its volume. According to Strevens (2008), one might say that

(a₂'') Gases behave as if there are no collisions of molecules.

This is asserted on the grounds that whether or not there are such collisions, the demonstration of Boyle's law goes through. It is clear that collisions are only negligible for the purposes of explaining Boyle's law. They are not negligible for other purposes (e.g., modelling diffusion). The attribution is that there are no collisions. Note that the falsity of this claim is not relevant in evaluating the truth of the as-if claim. We know that it is false, but what really matters is the argument given for the negligibility claim. Strevens cites McQuarrie and Simon (1997, p. 1015), who argue that 'any collision that deflects the path of a molecule from [the path assumed in the derivation] will be balanced by a collision that replaces the molecule.'

There are thus three kinds of negligibility claim. The first is that causal forces that are known to affect the behaviour of interest are weak enough to be negligible in strength (e.g., air pressure). Secondly, forces that are taken to be non-negligible in strength and potentially relevant to the behaviour in question have negligible net effects (due to some counterbalancing of forces, as with molecular collisions). Thirdly, any model ignores an infinite number of causal factors because those factors are thought to be irrelevant. In this last case, the as-if locution is not used. One does not say, for example, that 'stones fall as if swallows do not migrate to warmer territories during the autumn' because one presupposes that the migration behaviour of these birds could not affect the stones in any way.

There is also a difference between (a₁) and (a₂). A feather does not satisfy the relevant characteristics, and one should not say that 'the feather falls as if it were following Galileo's law' because *the as-if claim would be false*. Contrast this with (a₁). If the reason for using the as-if locution is indeed only to indicate that the truth of (a₁) does not matter, the falsity of the claim would not prevent the modeller from using the locution. He or she would use it even knowing that people do not optimize. The difference between (a₁) and (a₂) is that the former does not express a claim that is taken to be causally important for the model, but the latter does. In case a₂, the as-if locution is needed for specifying the exact content of the behavioural as-if claim, whereas in case a₁ it merely indicates that the behavioural claim is not to be evaluated for truth.

Mäki (1998, 2003) suggests that (AIC) could also be taken to express epistemic uncertainty concerning whether A is B. It seems that if this is applicable at all, it is most suitable for inaccurate as-if claims (case a₁). One would say that 'people behave as if they were rational' because one is uncertain whether they are rational. Note, however, that such use violates the pragmatic constraint for using the as-if locution. One cannot presuppose that the attribution is false if one is simultaneously uncertain about its truth.¹¹ One might argue that provisionally expressing epistemic uncertainty with respect to the attribution counts as not being committed to its truth and thus that the constraint is satisfied after all. But this will not do because, whoever asserts the aforementioned as-if claim in order to express epistemic uncertainty is ultimately committed to the truth of the attribution: if later investigations show that the attribution is false, he or she is no longer willing to assert the behavioural claim. That is what being committed to the truth of the attribution means.

Distinguishing between different kinds of as-if claims on the basis of whether their truth depends on the truth of the attribution is important for specifying the *object* of epistemic uncertainty. It may concern the attribution, the as-if claim and the explanation of whether the claim may be true, and if so, why. In cases (a₁) and (a₂) epistemic uncertainty

concerning the attribution *implies* that of the as-if claim. Because Mäki does not distinguish between as-if claims and attributions, he does not distinguish between different objects of uncertainty. He is thus led to conflate epistemic uncertainty concerning ‘why it is that A behaves in the way it does’ with uncertainty concerning ‘whether A is, or is not, B’ (Mäki, 1998, p. 26). I suggest that as-if locutions are not used for expressing epistemic uncertainty precisely because epistemic uncertainty concerning the attribution does not imply epistemic uncertainty concerning the as-if claim in the most genuine cases (b and c). Using it for such purposes would perpetuate the false impression that the truth of an attribution is always relevant to the truth of the as-if claim.¹²

As-if claims never explain ‘why it is that A behaves the way it does’ *in themselves*. This feature is difficult to reconcile with realism, even concerning case (a₂). Indeed, the use of as-if locutions can be justified because they give an indication that the claim itself does not provide an explanation.

4. False ascription as-if claims

The second use of the as-if locution involves ascribing false intentions or cognitions to something in order to describe its behaviour. The basic form of such a claim is thus the following:

(b) A behaves as if it had intentions or cognitions to B.

The point is to *describe behaviour* accurately when there is no need to concern oneself about whether the intentional or the cognitive attributions are correct. Note that in contrast to case (a₁), here one aims to assert that the behaviour formulated by means of the as-if claim is accurate.

Many of Friedman’s (1953, p. 21) examples exemplify such unrealistic attributions: one can describe the trajectories of billiard balls accurately under the assumption that expert players act as if they know complicated mathematical formulas derived from physical theory. This theory attributes the ability to make such complicated calculations to the players. The attribution is false because the players do not conceptualise the issue in terms of the physical theory. Nevertheless, one could derive trajectories for the billiard balls by using that physical theory. Similarly, Friedman (1953, p. 19) argues that the position of the leaves on a tree may be accurately described on the assumption that each individual leaf acts as if it is trying to maximise the amount of sunlight it receives. Lehtinen and Kuorikoski (2007) posit that in such cases the assumption is *behaviourally realistic* but *intentionally unrealistic* because it describes the behaviour truthfully but attributes intentions untruthfully. Ascribing intentions to the tree fails to accurately identify the causes of the leaf pattern.¹³

In Friedman’s examples of billiards players and leaves the attribution is clearly false, and the truth of the as-if claim does not depend on the truth of the attribution. Examining the truth of the attribution does not provide relevant information about the truth of the as-if claim. Friedman also uses this kind of as-if claim in arguing that firms behave as if they were maximising profits (1953, p. 21). He was perfectly aware of the fact that managers do not intentionally try to equalise marginal cost and benefit, and that they often do not even know the cost function. The point of using the as-if claim, however, is to argue that even if they do not intentionally try to maximise profits, their behaviour is so close to the profit-maximising assumption that one could obtain good predictions on that basis. These predictions do not, of course, concern profit maximising in itself, but rather focus on other things such as the effects of a price change or a tax on supply. Friedman was arguing that

the profit-maximisation assumption was, in fact, accurate for the purposes for which it was meant in standard price theory. I take it that such an interpretation is no longer considered radical, given that Hoover (2009) also argues that Friedman cares about the accuracy of assumptions.

If the attribution does not provide relevant information for evaluating the truth of this kind of as-if claim, what does? Friedman is known to rest his case on an evolutionary-selection argument: only firms whose behaviour mimics profit maximising will survive in the long run. The selection argument is commonly taken to facilitate explanation of why the behaviour of A could mimic that of B even when A is not B in various different contexts. Lagueux (1994) argues, however, that as-if claims are always justified by some more general theory, but he does not require this theory to be that of evolutionary selection. This seems right, given that there does not seem to be any need for a selection argument in the case of Galileo's law of falling bodies. There is a general theory, but unlike in Friedman's other examples (firms, leaves on a tree, billiards players), it is not separate from the as-if claim because it is the very theory of gravity that explains why one can describe objects as if they followed Galileo's formulation $s = \frac{1}{2}gt^2$.

The selection argument tends to be considered less than fully compelling (see Vromen, 2009 for an account and references), and therefore the profit-maximisation assumption is often considered inaccurate after all. Even Becker (1962, p. 10) argues that the assumption is not accurate.

In my judgment the great achievement of the "survival" argument advanced by Alchian and others' is not a demonstration that surviving firms must act as if they were trying to maximize profits, *for counterexamples can easily be developed*, but rather a demonstration that the decisions of irrational firms are limited by a budgetary constraint.

The fact that Friedman employs the survival argument gives support to the interpretation that Friedman *tried* to argue for the accuracy of the profit maximisation assumption and thus that his as-if claim concerning firms is to be interpreted in terms of (b) rather than (a₁). Yet, I would hesitate to argue that Boland's interpretation of Friedman in terms of (a₁) is wrong because Friedman thought that discussing the 'realism' of assumptions is only meaningful relative to the purpose of theorising. Thus, although he seems to have thought that profit maximising is accurate enough for the purposes of the theory of the firm as it was practised in the 1950s,¹⁴ he probably thought that discussing whether the assumption is accurate in an absolute sense is meaningless.

I do not wish to rely on a correct interpretation of Friedman. I will instead examine the use of the claim in another economic model in order to show that it could be used to describe behaviour accurately and even realistically. Thus, the failure of Friedman's example from economics should not lead one to believe that this kind of as-if claim is always inaccurate. Given that there is no need to provide any kind of selection argument or general theory in this case, it will also demonstrate that such arguments and theories are not *necessary* for the realistic description of behaviour through as-if claims.

4.1. A case study: signal extraction in an expected utility model of strategic voting

Consider Lehtinen's (2006, 2007, 2008) use of a signal extraction model as part of an expected utility model of strategic voting. Strategic voting is usually defined as giving one's vote to a candidate or alternative that is not considered to be the best. Only the case with pair-wise comparisons is analysed here in order to simplify the presentation. Lehtinen assumes that voters do not know other voters' preferences, but that they act as if they had obtained informative signals concerning them. Let N denote the total number of voters and

n_{jk} the number of voters who prefer alternative j to k . If $n_{jk} > N/2$, j will beat k in a pairwise comparison under the majority rule if everyone votes sincerely. The model is set in a computer-simulation framework in which voters' preferences are generated randomly. Voters are assumed to act as if they have obtained perturbed information on the realised value of n_{jk} as follows: an individual voter's signal $s_i(jk)$ is given by the sum of the standardised n_{jk} and a random term:

$$s_i(jk) = \frac{2n_{jk} - N}{\sqrt{N}} + \varepsilon R_i, \quad (1)$$

where R_i denotes a standard normal random variable and ε denotes a parameter that reflects the quality of the information. He shows (Lehtinen, 2006) that an individual's probability $p_i(jk)$ for the proposition that alternative j beats alternative k is given by

$$p_i(jk) = 1 - \Phi\left(-\frac{s_i(jk)}{\varepsilon\sqrt{1 + \varepsilon^2}}\right) + \varepsilon R_i, \quad (2)$$

where Φ denotes the standard normal distribution function. The existence of such signals is not a fiction because in mass elections pre-election polls may be interpreted as some sort of signal concerning the winning probabilities of various candidates, and even in parliamentary voting the representatives have various sources of information on other voters' preferences, such as party affiliation and the content of arguments made by different representatives. Yet, various features of the signals are fictions: the derivation in the appendix consists of four pages of manipulations of random variables and their densities. It would be foolhardy to claim that voters are able to make such calculations, let alone that they would actually perform them. Real voters do not engage in computing such expressions at all. Similarly, they do not literally compute expected utilities by multiplying probabilities by utilities. Furthermore, in the real world they do not obtain perturbed signals in the kind of form in which they are represented in the model, in which a signal is given by an exact numeral. For example, a signal could be given by $s_i(jk) = -0.298234$ and a probability by $p_i(jk) = 0.267543$.

These models of strategic voting thus depict voters as if they had obtained perturbed signals in an exact numerical form, as if they were able to perform complicated calculations and as if they actually carried out the computations and used them in an expected-utility calculus for determining their decisions. Now compare the following two statements:

- (b') Voters act as if they have obtained perturbed signals concerning the preference profile and as if they have computed probabilities from these signals using statistical reasoning.
- (b*) Voters obtain perturbed signals concerning the preference profile and compute probabilities from these signals using statistical reasoning.

(b*) is clearly false but (b') is true enough in the sense that actual voters' behaviour really corresponds to that which would ensue if they actually obtained signals, computed probabilities and so on. Introducing the as-if locution converts a false statement into a true one and the semantic change is obvious.

These strategic voting models are intentionally and cognitively unrealistic in the sense that voters have cognitions and intentions (signals and computing) in the model that they do not have in reality. However, the model may nevertheless be perfectly accurate in

accounting for their behaviour, and even realistic in the sense of correctly identifying the factors that influence the behaviour and the way in which they do so. I will now attempt to demonstrate this.

What would a realistic description of strategic voting behaviour look like? One should get the determinants of the behaviour right. Consider Anthony Downs' verbal description of the incentives for strategic voting:

Assume that there are three parties: Right, Center, and Left. Voter X prefers Right to Center and Center to Left, but he *believes* that Right has the least chance of winning. If he *greatly prefers* Right to Center and is *almost indifferent* between Center and Left, he is *less likely* to switch his vote from Right to Center than if he *slightly prefers* Right to Center but *abhors* Left (Downs, 1957, p. 49).

The implication here is that strategic voting depends on the winning chances of various candidates and on preference intensities. Let us say that voters engage in *Downsian deliberation* if they conceptualise voting in such (qualitative) terms. It involves the mental operation of weighing the desirabilities of the options with the probabilities. This operation could more generally be called *expected-utility deliberation*. Lehtinen's models are based on Enelow's (1981) and McKelvey and Ordeshook's (1972) expected utility models for explaining voter behaviour under various voting rules. These models show how beliefs and winning probabilities ought to combine so as to take intensities and probabilities into account in the right way.

If all one wishes to do is to provide an abstract account of people's strategic voting behaviour, such models are sufficient. However, in order to study the aggregate-level consequences of such behaviour it is necessary to give specific values to individual utility functions and beliefs, and the question that then arises is how they should be determined. Lehtinen goes on to generate utilities randomly and uses the signal extraction model to derive the beliefs. Given that behaviour depends on voters' beliefs, his model should generate beliefs in a realistic way.

Here it seems plausible to state that there are interpersonal differences in the *accuracy* of knowledge concerning the winning chances of various candidates and in the *confidence* different voters have in their knowledge. If one believes that the Right does not have a chance of winning but is wrong about this and votes strategically, and if the outcome then turns out to be one's second-best option rather than the best one due to that strategic vote, one would have hurt oneself by voting strategically (see Lehtinen, 2007 for an example). Hence, it seems reasonable to require that people vote strategically only if they can be reasonably confident that their information is correct. Furthermore, and even more obviously, the beliefs derived from the signal extraction model should be responsive to variations in the realised preference profile. This is precisely what the signal extraction model shows (see esp. Lehtinen, 2006 for details).

Even though voters must be said to act only as if they had received signals and as if they had computed beliefs on the basis of those signals, the signal extraction model provides a perfectly realistic account of their deliberation and the determinants of strategic voting. The degree of confidence in and the real quality of the information are the actual determinants of their beliefs when they vote strategically, and the model captures them in a realistic manner.¹⁵ It is thus realistic with respect to some intentions but unrealistic (and fictional) with respect to some others.

Note that whereas the attribution states that voters intentionally calculate these probabilities, the as-if claim merely states that the actions (or beliefs) derived from the model may be taken to describe actual voters' actions (or beliefs). Furthermore, it is definitely not assumed that voters have particularly good cognitive capacities. If ε is large,

some voters may have highly inaccurate beliefs. It would be silly to criticise this model for attributing unrealistically sophisticated behaviour to voters on the grounds that it describes their behaviour as if they had gone through complicated calculations. Using the as-if locution highlights this fact.

Formulating the behavioural assumption in statements such as voters act as if they computed probabilities on the basis of perturbed signals does not indicate whether or not the behaviour will be realistically modelled. The as-if claim in *itself* does not justify such thinking, it merely formulates the behavioural claim. Nevertheless, the fact that the attribution (b*) is unrealistic is entirely irrelevant in terms of evaluating the realism of the as-if claim: it is known that voters do not go through complex computations when they vote, but that is irrelevant. However, merely stating that (b') is true enough is not sufficient because it does not provide any explanation of why one should think that it is. Its truth is rather to be demonstrated. Hence, as-if claims are typically *incomplete* in that showing that the model is realistic always requires something beyond the claim itself. Other kinds of as-if claims also seem to be incomplete in this way.

As demonstrated above, whether negligibility-indicating (a₂) as-if claims involve making accurate assumptions depend on the degree to which the attribution is true. In this case (b) there is no similar continuum of claims from the perfectly accurate to the utterly inaccurate. The claim that voters obtain signals expressed in terms of a seven-digit number, compute probabilities on this basis and choose an action after considering the expected utility of the various options is a fiction in other words known to be false. However, I argue that the signal extraction model does accurately capture the determinants of behaviour.

There is no general theory (such as evolutionary selection) that explains why individual beliefs should correspond to those generated in the model. I have argued that the fiction of computing behaviour and its signals allows for a realistic representation of voting behaviour. Ultimately, I assert that the as-if claim is realistic on the grounds that it correctly describes Downsian deliberation, which, I also assert, provides a true account of the reasons for strategic voting.

Morrison (2009) claims that the way in which fictional representations provide reliable information differs from one case to another. It does not seem possible to generalise on why, whether and when they facilitate the realistic description of behaviour, however, which could be attributable to the fact that fictions play a crucial role in *representing* the world (see also Teller, 2009). In this case it is necessary to represent the signals, the utilities and the probabilities with precise numerals in order to represent Downsian deliberation in a computer model, and the function of the as-if locution is to show that this representation is not intended to provide a literally true description of all the details of that deliberation (cf. Coddington, 1979).

4.2. A digression: Mäki's distinction

Mäki (1992, 1998, 2000, 2003, 2004, 2009) distinguishes (a₂).

- (a₂) Phenomena behave as if certain ideal conditions were met: conditions under which only the real forces that are theoretically isolated are active

from (b[~]) thus:

- (b[~]) Phenomena behave as if those forces were real.

I will now attempt to clarify how the classification proposed here differs from Mäki's distinction. Mäki's distinction concerns the *reference* of the as-if locution: he

distinguishes between cases in which something is modelled as if certain minor forces were absent, and those in which the central forces in the model are based on fictions. In the former case the isolation of forces is false (i.e., they are not the only forces that operate in reality), whereas in the latter case the central forces picked out are fictional (i.e., they are found in the representation but not in the real world) (Mäki, 2009). This distinction cannot be incorporated into the classification scheme proposed here because the basis on which it rests is different. Mäki's distinction is based on what the fiction the as-if locution expresses is about, whereas here it is based on the pragmatic reason for using the locution. Mäki does not give any reason why one would want to model the central causal forces with a fiction.¹⁶

According to Mäki (1998), (b^{\sim}) invites an instrumentalist reading because it postulates fictional forces. In my example of signal extraction individual behaviour is described in terms of fictions even though it is a central causal factor in the model. Mäki's distinction thus requires that (b') must be understood in terms of (b^{\sim}), and (b^{\sim}) should thus be understood in instrumental terms. I believe this conclusion should be resisted.

I have argued in this section that an as-if claim may well be realistic even in a case in which central forces are represented by fictions. If I have succeeded in showing that voters *really* behave as though they were governed by some fictional forces (computing probabilities and expected utilities) identified in the model, there seems to be little reason for a realist to be worried. The relevant question is how the realistically described behaviour specified in the as-if claim is justified or explained given that the claim itself does not provide an explanation. It is obvious that such an explanation must be found elsewhere, and what I have written in this section provides it. The difference between realists and instrumentalists is presumably that the former require that an explanation or justification is provided somewhere, whereas the latter do not care as long as the predictions provided by the theory are correct. Only a foolish brand of realism would require that each and every assumption must be formulated in such a way that an explanation is also provided in the very formulation. As demonstrated above, the attribution is plainly false, and the as-if locution is needed in order to effect a semantic change that allows the correct understanding of the behavioural assumption. If realists can accept unrealistic assumptions if they are properly justified, they should not banish realistic assumptions merely because they are represented by fictions.

Mäki could perhaps take a different tack, and argue that my example does not count as a case of (b^{\sim}) because, although obtaining signals and multiplying utilities are central forces in the model, they are real. It is true that the as-if claim does not qualify the *existence* of these forces. However, neither does it indicate isolation or negligibility. It does not mean that there might be other relevant forces that are ignored for the purposes of investigation. The as-if locution is rather needed to indicate that the modeller *does not claim* that voters compute probabilities or multiply them by utilities, albeit the fictions of literally manipulating density functions and of literally multiplying utilities by beliefs are needed for formulating the behavioural assumption. It would be misleading to assert that voters compute probabilities or multiply them by utilities. These fictions are best compared to using physical theory to determine the trajectories of billiard balls in Friedman's example: it is known that these fictions are not the underlying causes of the phenomena of interest. If Mäki were to adopt this second line of reasoning, he would end up in a dilemma: he would either have to explain why using physical theory does not provide a fictional account of billiards players' intentions or show how Friedman's example is relevantly different from my signal-extraction example.

5. Non-committal-indicating as-if claims

Dennett (1996, p. 27) suggests that the intentional stance is the strategy of interpreting the behaviour of an entity (person, animal, artefact) by treating it *as if* it were a rational agent who governed its 'choice' of 'action' through 'consideration' of its 'beliefs' and 'desires'. Here the as-if claim expresses a non-committal attitude towards whether the entity in question has beliefs and desires. Note that even though the as-if locution is used to bracket the existence of beliefs and desires, the assertion about behaviour is intended to be accurate. This is what makes Dennett's use of the as-if locution different from that of inaccurate as-if claims (case a_1). Given that a thermostat really behaves as if it wanted to keep the temperature fixed, such as-if claims may be used to describe behaviour accurately even if the attribution (thermostats have desires) were to be blatantly false. What makes the behavioural claim about thermostats true is that they have been designed to do just that: keep the temperature fixed.

When the intentional stance is applied to real people, the relevant attribution is that 'people have desires and beliefs'. For someone persuaded that people have desires and beliefs, this would seem to provide a counterexample to the thesis that attributions are always false. However, Dennett would use the as-if locution even if his audience thought that people had beliefs and desires. He would say that this attribution cannot be proven to be true, and that the as-if locution is used precisely to indicate that the existence of desires and beliefs is bracketed. Thus it is ultimately not the falsity of the attribution that pragmatically justifies using the as-if locution but rather whether the utterer of the locution is committed to its truth.

Let us now consider an example from economics which is more difficult to interpret.

- (c) *Orthodox revealed preference theory* then provides consistency conditions for Eve's behavior in A to be described by saying that she chooses *as though* maximizing the expected value of a utility function defined on C, relative to a subjective probability distribution defined on B (Binmore, 1998, pp. 360–361).

In order to see the relevant attribution for this claim, it is necessary to look at the notion of utility more closely. Classical utilitarians used the term 'utility' to denote either a mental state (degree of satisfaction) or a feeling (e.g., pleasure). 'Maximising utility' could then be understood as an intentional activity whereby an agent endeavours to maximise the degree to which his or her desires are satisfied. If 'utility' were to be interpreted in such a manner, one would say that the attribution is that 'Eve maximises desire satisfaction', which would be true if she actually maximised her desire satisfaction. Note that if Eve actually maximised utility in this sense, such an interpretation would constitute a counterexample to the claim that the attribution is always false.

Here, however, Binmore would use the as-if locution in describing Eve's actions even if Eve deliberately and consciously endeavoured to maximise satisfaction. By way of illustration, consider the way in which 'utility' and 'maximisation' are interpreted in modern utility theory, which essentially consists of *representation theorems*.¹⁷ Such theorems specify that if preferences satisfy a set of *consistency conditions*, then they may be represented with utility functions. Let P_i denote individual i 's strict preference ordering. The standard interpretation is that for some alternatives x and y , xP_iy means that i considers x better than y . The representation theorem then states that $U_i(x) > U_i(y)$ if and only if xP_iy . In other words, function U assigns a higher number to x than to y if the individual strictly prefers x to y . It could then be said that the individual's behaviour is described as if he or she were maximising a utility function.

The theory features utility functions that may be used to describe the behaviour of an individual whose preferences satisfy the conditions. Sometimes the term ‘utility’ or ‘util’ is used for the values of these functions. Given that ‘utility’ refers to a number here, it consists neither in material welfare nor in any kind of mental state such as pleasure. Binmore would be making a nonsensical statement that involved a category mistake if he were to drop the as-if locution. It is nonsensical to say that ‘a person tries to maximise utility’ when utility is interpreted in its modern sense as a description of the person’s preferences or choices. Maximisation is not a mental operation carried out by the person whose behaviour is described, it is rather a property of a mathematical model that represents that person’s behaviour in terms of maximisation. The category mistake would thus consist in stating that ‘a person aims to (purposefully) maximise his or her utility function’. However, the person’s actions could be described as if he or she were maximising a utility function because the more preferred (or chosen) outcomes are assigned higher utilities than the less preferred. As Binmore states, ‘our story makes it *nonsense* to say that Adam chooses the former *because* its utility is greater’ (Binmore, 2009, pp. 19–21). Saying so would amount to committing the ‘causal utility fallacy’. In other words, utility does not provide any reasons for choosing one way or the other, and merely represents an individual’s preferences.

Binmore could not use the locution in order to claim that actual behaviour corresponded to that which would ensue if Eve maximised her desire satisfaction because he claims that modern decision theory ‘makes a virtue of assuming nothing whatever about the psychological causes of our choice behavior’ (Binmore, 2009, pp. 8–9). Representation theorems do not say anything about why people are supposed to maximise utility. Unlike in case (b), Binmore does not attribute false intentions to Eve in order to describe her behaviour accurately.

Neither does Binmore qualify the behavioural claim with an as-if locution because he denies that Eve maximises utility (in the modern sense). He has already assumed that Eve is rational:

Although the revealed preference interpretation of utility will be maintained throughout this book, half of the labor of constructing a utility function from an agent’s choice behavior will be skipped. It will be assumed that a preference relation has already been constructed and that it remains only to show that it can be represented using an appropriate utility function (Binmore, 1994, p. 268).

Even though he readily accepts elsewhere that ‘human behavior is often downright irrational’ (Binmore, 2005, p. 75), he cannot doubt Eve’s rationality here because preference relations cannot be constructed at all unless individuals are rational. The locution is thus used in order to formulate an accurate behavioural assumption.

One might propose that maximising expected utility is to be interpreted as expected utility deliberation. ‘Eve maximises expected utility’ would then be the relevant attribution, and maximisation would be interpreted in terms of this mental operation. This would not be correct either, however, because the criterion of truth for the as-if claim ‘Eve’s behaviour can be described as if she were maximising the expected value of a utility function’ which is not whether her behaviour corresponds to that which would ensue if she engaged in expected-utility deliberation. It is rather whether her preferences satisfy the conditions for representing utility. Whether a person’s actions could be described as if he or she were maximising expected utility does not depend on whether he or she engages in some kind of mental operation. Similarly, whether or not a model incorporating the expected utility assumption is accurate does not depend on whether the person is trying to maximise something or whether he or she is engaging in expected-utility deliberation, but

only on whether his or her preferences satisfy the conditions. I take this to be the standard position among decision theorists, but I will provide a simple argument in favour of it: Eve may violate the conditions even if she is engaged in expected-utility deliberation if, for example, she makes errors in calculating expected utilities.

In the previous section the attribution was:

(b*) Voters obtain perturbed signals concerning the preference profile and compute probabilities from these signals using statistical reasoning.

A more fully specified attribution would have been:

(b**) Voters obtain perturbed signals, compute probabilities and literally calculate expected utilities by weighing utility values with probabilities.

Literally calculating expected utilities by multiplying probabilities and utilities expressed in seven-digit numbers is thus a mental operation that is also falsely attributed to voters. Given that (b) and (c) both concern cases with expected utility maximisation, one might argue that they are not different enough to justify putting them into separate categories. Lehtinen (2011) posits, however, that the credibility of his model crucially depends on the plausibility of the assumption that voters really engage in Downsian deliberation, and that whether their preferences satisfy the consistency conditions matters very little in the sense that the results of the model do not depend on it. In fact, it seems that satisfying the consistency conditions is not sufficient to guarantee that voters engage in Downsian deliberation. This is at least what one is compelled to think if one accepts Binmore's argument that modern utility theory makes no psychological assumptions.

Completely different things are required in the two cases for the as-if claims concerning expected utility maximisation to be true. In case (c), the claim is true if the individual satisfies the consistency conditions. In case (b), however, whether or not it is true enough depends on whether or not voters engage in Downsian deliberation. The similarity between (b) and (c) lies in the fact that, in both cases, the modeller uses an as-if locution in order to clarify what kind of claims about behaviour and mental states are made. Binmore uses the as-if locution to alert the reader to the fact that 'maximisation' does not refer to any kind of mental operation. He avoids making the false claim that maximisation is such an operation by using the locution.

Let us now return to (c). Once the right attribution is specified, it is clear that it is indeed always either false or involves a category mistake. I suggest two possibilities. The first is that it concerns Eve's intentions: 'Eve endeavours to maximise expected utility' when utility is understood in its modern sense as a number describing preferences. This attribution incorporates a category mistake. It is not a consequence of an empirical failure to act in a certain manner but rather derives from the fact that utility is not something that Eve can try to maximise as a matter of the logical form of the term 'utility'. The second possibility is that it concerns Binmore's beliefs concerning his audience. He could be taken to assume that not everyone in his audience knows that the term 'utility' will be used in its modern sense. He would thus be using the as-if locution because he believes that some people in his audience might think that 'utility means pleasure in the theory that Binmore is presenting'. This is again false. It is also evident that evaluating the truth of either of these attributions is entirely irrelevant to whether or not the as-if claim is accurate because neither says anything about the consistency conditions.

One might argue that if the first attribution is appropriate, claims like (c) should be treated as a separate *fourth* kind of as-if claim on the grounds that avoiding category mistakes is not intrinsically related to expressing non-committal to any particular

determinants of behaviour. In further support of such an argument one could point out that the above quotation from 1994 commits Binmore to Eve's rationality, and that rationality is a determinant of behaviour. I resist such an interpretation mainly because rationality is not the right kind of determinant. Insofar as the 'preference relation has already been constructed', the behaviour has also already taken place, and if it is described in terms of utility functions, it does not matter what its determinants were. Furthermore, once one understands why Binmore uses the as-if locution, one also understands how avoiding category mistakes amounts to being non-committal about the determinants of behaviour.

As-if claims used in expected utility theory may well provide accurate accounts of behaviour. However, here the use of the as-if locution tells us absolutely nothing about whether or not this is the case. It has nothing to do with the question of whether or not it is accurate to describe people's actual behaviour as utility-maximising. This question can be evaluated in many different ways, of which one is to evaluate whether people in some actual situations try to maximise pleasure. This is so because the reasonability of the consistency conditions depends on how real people act, and one of the relevant considerations in assessing whether their behaviour satisfies the conditions is whether or not they try to maximise pleasure. In some circumstances they do satisfy the conditions, in others they do not, but in any case one needs to use the as-if locution even when expected utility maximisation is an accurate assumption. This is what makes (c) different from case (a₁). If one wished to say that the difference between actual behaviour and that specified by expected utility theory is negligible, one would have to say that 'people act as if they were acting as if they maximised expected utility'. Naturally, such claims are never made because there are more elegant ways of conveying the idea. One could say, for example, that 'Eve acts as if she were a utility maximiser'.¹⁸ This is because the relevant attribution 'Eve is a utility maximiser' is false if she does not act according to the conditions.

Some decision theorists, including Binmore, apparently think that mere choices rather than preferences ultimately need to be consistent for the theory of expected utility to apply. One could then state that 'an agent acts as if he or she had consistent preferences', and that his or her choices could be described in terms of utility functions. Here the very existence of preferences is bracketed. Note that if expected utility theory is interpreted in such revealed-preference terms, stating that 'an agent acts as if he or she had consistent choices' would mean that the as-if claim is to be interpreted as indicating inaccuracy (a₁).

In cases (b) and (c) the reason for using the as-if locution is that it allows one to specify exactly what one wants to say about behavioural assumptions. Dennett and Binmore use it to emphasise that no intentional ascriptions need to be made, whereas Lehtinen uses it to formulate a realistic behavioural assumption and to show that the credibility of the model does not hinge on the assumption that voters actually engage in computing expected utilities or receive numerically precise signals. These cases are similar in that using the as-if locution allows the modeller to *avoid* making some *particular* mental attributions.

6. Conclusions

As-if locutions can be used in a variety of ways. I have argued that they can be used to (a₁) indicate that the behavioural claim they qualify is inaccurate but negligibly so, and (a₂) to indicate that the modeller acknowledges the existence of some causal factors that are not explicitly taken into account in the model because they are deemed to be negligible. In the most genuine cases, they are used for (b) expressing the content of accurate behavioural claims by ascribing intentions on entities in an unrealistic manner, and (c) for indicating that descriptive claims made about behaviour do not commit one to any particular mental

assumptions. The claims could thus be put briefly as follows.

- (a₁) *Inaccuracy-indicating claims.* AIC – or rather, A does not actually behave as B, but that is unimportant in this model.
- (a₂) *Negligibility-indicating accurate claims.* AIC because A is like B in some respects but not in all. The ignored respects are not important in this model.
- (b) *Claims indicating false mental ascription.* AIC because A does not have (mental) properties B and does have properties C. Yet its behaviour could be described under the assumption that it has properties B.
- (c) *Claims indicating non-committal.* AIC because, although A has nothing to do with B, A behaves as B whatever may be the actual determinants of A's behaviour.

However, as-if claims also have some commonalities. First, their most important function is to help in specifying which part of a model ought to be true and what can be allowed to be false. Secondly, they would appear to be invariably related to making claims about or describing behaviour. Thirdly, they all express lack of commitment to the truth of their underlying attributions. Fourthly, in themselves they are always methodologically incomplete in the sense that they do not provide any explanation of why behaviour should, in reality, correspond to that specified in the model.

As-if claims are likely to create a great deal of confusion. In view of their incompleteness it is always necessary to find out whether a given one is realistic. Furthermore, it may be difficult to see the difference between the different kinds of claim. One could easily be misled into believing that an as-if claim is realistic even when behaviour is not realistically modelled, or that all as-if claims are false because their attributions are always false. Friedman uses the locution for all of the reasons described in this paper. Given that the different uses are based on different methodological commitments and lead to different ways of evaluating the truth of the claims, it is not clear that it makes any sense to talk about 'Friedman's as-if methodology'.¹⁹

Some people may prefer to minimise the use of as-if claims just because they are so confusing and incomplete. In all the cases in which they involve a semantic change, they clearly have a useful function. Usually the change involves transforming a false attribution into a true enough as-if claim. Yet, they are never necessary. If one merely wishes to express the idea that the realisticness of an assumption does not matter, or that one is aware of making an unrealistic assumption, I suggest that it would be better to avoid using the as-if locution because there are more natural and less confusing ways of making such claims. Whether a semantic change thus entailed could then be taken to facilitate recognition of the warranted uses of as-if claims.

This paper may be considered somewhat odd because the author is using his own work as an economist in a philosophical case study. However, the motivation for writing this paper arose from the experience of writing the papers on voting and signal extraction. Economic journals do not allow authors to explain the as-if justifications of their assumptions at length. The 'as-if' locution actually never appears in any of my papers on voting and signal extraction. It does not appear because I was too afraid that I would be interpreted as promoting my model in terms of inaccuracy-indicating as-if claims. This interpretation seemed to be very common among the economists, and I was so deeply concerned about being misunderstood that I decided not to use the locution at all. These experiences also explain why I have not been content merely to describe how as-if locutions are used in economics, but rather also tried to delimit their use with the pragmatic constraints: I believe that a more regimented use of the locution would help in avoiding confusions.

Given that (a₂), (b) and (c) represent cases in which the as-if claim may be accurate or even realistic, restricting usage to these three kinds also encourage evaluation of the criteria as to their truth. The most important consideration here is whether or not the truth of the attribution is relevant to the truth of the as-if claim. The example of signal extraction was intended to show that in one case in which it is not relevant, i.e. when it involves a fiction of a false intentional ascription, it may well be realistic despite being based on a fiction, and the grounds for evaluating its truth need not come from a more general theory.

Acknowledgements

This paper has been presented in the INEM conference in Xalapa, Mexico, and in the TINT seminar in Helsinki, Finland. Till Grüne-Yanoff, Jaakko Kuorikoski, Uskali Mäki, Luis Mireles-Flores, Samuli Pöyhönen and anonymous reviewers provided useful comments on earlier versions of the paper. The usual disclaimer applies.

Notes

1. The methodological argument that could justify this behavioural claim is attributable to Mill (2000 [1844], p. 98), but he never used the ‘as-if’ locution to express it.
2. For the latest discussion on Friedman, see *The methodology of positive economics: Reflections on the Milton Friedman legacy* (ed. Uskali Mäki). Vaehinger’s as-if account is discussed in Fine (1993), Mäki (2004), and Bokulich (2009), for example.
3. I am not arguing for this claim here. The reader is encouraged to find a case in which the as-if locution is necessary for formulating a claim. I have not found any such cases.
4. In most cases it is reasonable to interpret A as an object and B as a property that the object could have. If only objects can behave in certain ways, A must be an object. However, B need not always be a property of that object, as I show in the section on negligibility-indicating as-if claims. Consider, for example, the claim that ‘objects fall as if there were no air resistance’. The property of air resistance is relational because it requires reference to the surrounding circumstances (atmospheric pressure) in addition to the object itself. The formulation of AIC is deliberately silent on the interpretation of the terms in order to accommodate various different cases.
5. I add a caveat to this statement in the section on inaccuracy-indicating (a₁) as-if claims.
6. The reason for using attribution terminology is that property B is attributed to A. One might also say that ‘B is predicated about A’, or that ‘A is identified as being B’. However, using the term ‘attribution’ seems more natural here.
7. According to Friedman (1953, pp. 36–38), one can treat the same firm as if it were a perfect competitor on one question, and as if it were a monopoly on another. This example qualifies as an inaccuracy-indicating claim. Note that Kirman does not endeavour to justify the optimisation assumption, but rather ascribes the as-if justification to those who do. In fact, except for Friedman, I have not found examples of economic texts in which the inaccuracy-indicating as-if claim is made by the author who uses the corresponding assumption in a model. In contrast, MacDonald (2003) as well as Boylan and O’Gorman (1995) provide an example in which the inaccuracy-indicating as-if claim is attributed to others.
8. Another way of interpreting (a₁) is to say that it means ‘people do *not* optimize, but that does not matter for the model results’. If this were to be the interpretation, there would be a semantic change after all. The as-if claim ‘people behave as if they optimize’ would then mean that the attribution ‘People optimize’ is false (but negligibly so). My intuitions are slightly in favour of the interpretation provided in the main text, but I am willing to be persuaded to think otherwise. Adopting this interpretation would make analysing as-if claims more uniform with respect to semantic change but less uniform with respect to meaning. One would always have to look for a semantic change, but the meaning of the as-if claims would be the same only in cases a₂, b and c.
9. Friedman (1953, p. 40) also provides an example of this kind of as-if claim. Becker (1962, p. 5) writes ‘Impulsive households are assumed to act “as if” they only consulted a probability mechanism.’ It is not clear, however, what he means because he had already stated ‘Therefore,

- households can be said to behave not only “as if” they were rational but also “as if” they were irrational (p. 4).’ Becker’s lack of precision makes me wonder whether he intended to present an argument for rational choice theory (cf. Moscati & Tubaro, 2011).
10. Note that Friedman uses the term ‘assumption’ to mean what is called a ‘claim’ in this paper. If assumptions express propositions in models, one would say that ‘the model assumes a vacuum but the modeller does not claim that we live in a vacuum’.
 11. Note that Vaihinger distinguishes between *fictions* and *hypotheses* on the grounds that the former are known to be false whereas the latter are a matter of investigation.
 12. Naturally, I agree with Mäki that epistemic uncertainty is consistent with realism. I am merely concerned to show that expressing such uncertainty by means of as-if locutions does not satisfy my pragmatic constraints. In this sense, it is not surprising that Mäki has not provided any concrete examples of such uses.
 13. An anonymous reviewer suggested describing this example in terms of incorrect causes. It might be possible to generalise case (b) such that it concerns all false ascriptions of causes: ‘A behaves as if it were governed by causal forces B’. This is clearly possible at least in principle, but given that I have only seen examples of (b) in which the false attribution concerns mental states, I hesitate to argue for this more general formulation. Similar considerations apply to case (c).
 14. This aspect is even more evident in Machlup (1946).
 15. Lehtinen’s models are not intended to be realistic with respect to all aspects of behaviour. One major feature is that in his analyses of strategic behaviour he compares the case in which every voter engages in such behaviour to one in which nobody does so. What is presented here applies to voters *insofar as* they engage in strategic voting. In real life some voters engage in strategic behaviour and some do not. He relaxes this assumption, however, in Lehtinen (2010), but here is another sense in which the signal extraction model is not entirely realistic. It has been argued that voters systematically overestimate the winning chances of candidates they support, but the model implicitly assumes non-biased estimates. Note, however, that such biases could be taken into account, and if they were it would be via incorporating some parameters or functions into the formal model, and one would still have to use the as-if locution in presenting the computing part.
 16. The instrumental usefulness of fictions could be taken as a reason why as-if locutions are used in case (b*). Given, however, that stating that a fiction is instrumentally useful does not really explain why it is, it is rather begging the question, and it might be better to state that Mäki does not provide any reason for using the as-if locution in this case. It is because instrumental usefulness is question-begging that it is not included as a separate type of as-if claim, even though Friedman sometimes uses the as-if locution to indicate that a hypothesis is useful if it generates correct predictions (Friedman & Savage, 1948, p. 298, 1952, p. 473).
 17. The original contribution of this section is limited to the role of as-if locutions in expected utility theory. The ideas presented here were first clearly formulated by Luce and Raiffa (1957), and similar arguments are presented in Kavka (1991) and Lehtinen and Kuorikoski (2007).
 18. ‘Eve is a maximiser’ here is presupposed to mean that she satisfies the vNM conditions.
 19. I, too, have been guilty of using this expression (Lehtinen & Kuorikoski, 2007, p. 123).

References

- Becker, G. S. (1962). Irrational behavior and economic theory. *Journal of Political Economy*, 70(1), 1–13.
- Binmore, K. (1994). *Game theory and the social contract: Playing fair*, (Vol. 1). Cambridge, MA: MIT Press.
- Binmore, K. (1998). *Game theory and the social contract: Just playing*, (Vol. 2). London: The MIT Press.
- Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.
- Binmore, K. (2009). *Rational decisions*. Princeton: Princeton University Press.
- Bokulich, A. (2009). Explanatory fictions. In M. Suárez (Ed.), *Fictions in science: Philosophical essays on modeling and idealization* (pp. 91–109). London: Routledge.
- Boland, L. A. (1979). A critique of Friedman’s critics. *Journal of Economic Literature*, 17(2), 503–522.
- Boylan, T. A., & O’Gorman, P. F. (1995). *Beyond rhetoric and realism in economics: Towards a reformulation of economic methodology*. London: Routledge.

- Coddington, A. (1979). Friedman's contribution to methodological controversy. *British Review of Economic Issues*, 2(4), 1–13.
- Dennett, D. C. (1996). *Kinds of minds: Towards an understanding of consciousness*. London: Weidenfeld & Nicolson.
- Downs, A. (1957). *An economic theory of democracy*. New York: Harper.
- Elgin, C. Z. (2004). True enough. *Philosophical Issues*, 14(1), 113–131.
- Elgin, C. Z. (2009). Exemplification, idealization, and scientific understanding. In M. Suárez (Ed.), *Fictions in science: Philosophical essays on modeling and idealization* (pp. 77–90). London: Routledge.
- Enelow, J. M. (1981). Saving amendments, killer amendments, and an expected utility theory of sophisticated voting. *Journal of Politics*, 43(4), 1062–1089.
- Fine, A. (1993). Fictionalism. *Midwest Studies in Philosophy*, XVIII, 1–18.
- Friedman, M. (1953). The methodology of positive economics. *Essays in positive economics* (pp. 3–43). Chicago: University of Chicago Press.
- Friedman, M., & Savage, L. J. (1952). The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy*, 60(6), 463–474.
- Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, 56(4), 279–304.
- Hindriks, F. A. (2006). Tractability assumptions and the Musgrave-Mäki typology. *Journal of Economic Methodology*, 13(4), 401–423.
- Hoover, K. D. (2009). Milton Friedman's stance: The methodology of causal realism. In U. Mäki (Ed.), *The methodology of positive economics: Reflections on the Milton Friedman legacy* (pp. 303–320). Cambridge: Cambridge University Press.
- Kavka, G. (1991). Rational maximizing in economic theories of politics. In K. R. Monroe (Ed.), *The economic approach to politics: A critical reassessment of the theory of rational action* (pp. 371–385). New York: HarperCollins.
- Kirman, A. P. (2006). Demand theory and general equilibrium: From explanation to introspection, a journey down the wrong road. *History of Political Economy*, 38, 246–280.
- Kuorikoski, J., Lehtinen, A., & Marchionni, C. (2010). Economic modelling as robustness analysis. *British Journal for the Philosophy of Science*, 61(3), 541–567.
- Lagueux, M. (1994). Friedman's instrumentalism and constructive empiricism in economics. *Theory and Decision*, 37(2), 147–174.
- Lehtinen, A. (2006). Signal extraction for simulated games with a large number of players. *Computational Statistics and Data Analysis*, 50, 2495–2507.
- Lehtinen, A. (2007). The welfare consequences of strategic voting in two commonly used parliamentary agendas. *Theory and Decision*, 63(1), 1–40.
- Lehtinen, A. (2008). The welfare consequences of strategic behaviour under approval and plurality voting. *European Journal of Political Economy*, 24, 688–704.
- Lehtinen, A. (2010). Behavioural heterogeneity under approval and plurality voting. In J. Laslier & M. R. Sanver (Eds.), *Handbook on approval voting* (pp. 285–310). Heidelberg: Springer.
- Lehtinen, A. (2011). A welfarist critique of social choice theory. *Journal of Theoretical Politics*, 23(3), 359–381.
- Lehtinen, A., & Kuorikoski, J. (2007). Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences*, 37(2), 115–138.
- Luce, D. R., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.
- MacDonald, P. K. (2003). Useful fiction or miracle maker: The competing epistemological foundations of rational choice theory. *American Political Science Review*, 97(4), 551–565.
- Machlup, F. (1946). Marginal analysis and empirical research. *The American Economic Review*, 36(4), 519–554.
- Mäki, U. (1992). Friedman and realism. *Research in the History of Economic Thought and Methodology*, 10, 171–195.
- Mäki, U. (1998). As if. In J. B. Davis, D. W. Hands & U. Mäki (Eds.), *The handbook of economic methodology* (pp. 25–27). Cheltenham: Edward Elgar.
- Mäki, U. (2000). Kinds of assumptions and their truth: Shaking an untwisted F-twist. *Kyklos*, 53(3), 317–335.
- Mäki, U. (2003). 'The methodology of positive economics' (1953) does not give us the methodology of positive economics. *Journal of Economic Methodology*, 10(4), 495–505.

- Mäki, U. (2004). Realism and the nature of theory: A lesson from J H Von Thünen for economists and geographers. *Environment & Planning A*, 36(10), 1719–1736.
- Mäki, U. (2009). Unrealistic assumptions and unnecessary confusions: Rereading and rewriting F53 as a realist statement. In U. Mäki (Ed.), *The methodology of positive economics. Reflections on the Milton Friedman legacy* (pp. 90–116). Cambridge: Cambridge University Press.
- McKelvey, R. D., & Ordeshook, P. C. (1972). A general theory of the calculus of voting. In J. F. Herndon & J. L. Bernd (Eds.), *Mathematical applications in political science*, (Vol. IV), (pp. 32–78). Charlottesville: The University Press of Virginia.
- McQuarrie, D. A., & Simon, J. D. (1997). *Physical chemistry: A molecular approach*. Sausalito: University Science Books.
- Mill, J. S. (2000 [1844]). *Essays on some unsettled questions of political economy* (2nd ed.). Kitchener: Batoche Books.
- Morrison, M. (2009). Fictions, representations, and reality. In M. Suárez (Ed.), *Fictions in science: Philosophical essays on modeling and idealization* (pp. 110–135). London: Routledge.
- Moscatti, I., & Tubaro, P. (2011). Becker random behavior and the as-if defense of rational choice theory in demand analysis. *Journal of Economic Methodology*, 18(2), 107–128.
- Musgrave, A. (1981). ‘Unreal assumptions’ in economic theory: The F-twist untwisted. *Kyklos*, 34(3), 377–387.
- Røgeberg, O., & Nordberg, M. (2005). A defence of absurd theories in economics. *Journal of Economic Methodology*, 12(4), 543–562.
- Satz, D., & Ferejohn, J. A. (1994). Rational choice and social theory. *Journal of Philosophy*, 91(2), 71–87.
- Sen, A. K. (1973). Behaviour and the concept of preference. *Economica*, 40(159), 241–259.
- Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.
- Teller, P. (2009). Fictions, fictionalization, and truth in science. In M. Suárez (Ed.), *Fictions in science: Philosophical essays on modeling and idealization* (pp. 235–247). London: Routledge.
- Vromen, J. J. (2009). Friedman’s selection argument revisited. In U. Mäki (Ed.), *The methodology of economics. Reflections on the Milton Friedman legacy* (pp. 257–284). Cambridge: Cambridge University Press.