International Journal of Approximate Reasoning 180 (2025) 109388

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Assessing inference to the best explanation posteriors for the estimation of economic agent-based models

Francesco De Pretis [a,b], Aldo Glielmo [c], Jürgen Landes [d,e,*]

[a] Department of Communication and Economics, University of Modena and Reggio Emilia, 42121 Reggio Emilia, Italy
[b] Department of Environmental and Occupational Health, School of Public Health, Indiana University Bloomington, 47405 Bloomington, IN, USA
[c] Applied Research Team, Directorate General for Information Technology, Bank of Italy, 00044 Rome, Italy
[d] Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich, Germany
[e] Open Science Center, LMU Munich, 80539 Munich, Germany

ARTICLE INFO

Keywords:
Agent-based models
Asset pricing
Bayesian updating
Explanation
Inference to the best explanation

ABSTRACT

Explanatory relationships between data and hypotheses have been suggested to play a role in the formation of posterior probabilities. This suggestion was tested in a toy environment and supported by simulations by David H. Glass. We here put forward a variety of inference to the best explanation approaches for determining posterior probabilities by intertwining Bayesian and inference to the best explanation approaches. We then simulate their performances for the estimation of parameters in the Brock and Hommes agent-based model for asset pricing in finance. We find that performances depend on circumstances and also on the evaluation metric. However, most of the time our suggested approaches outperform the Bayesian approach.

1. Introduction

This paper develops techniques for parameter estimation based on a well-known philosophical approach, Inference to the Best Explanation (IBE) [1], that has attracted recent interest [2–13]. Our starting hypothesis is that *explanations* may do important confirmatory work beyond what can be covered by (Bayesian) probabilities [14].

While there is a debate on what exactly IBE is [14,15], the main idea is clearly that we ought to *infer the hypothesis that best explains the data*, since good explanations are truth-conducive [16], see further Section 2.1.1. The question arises of how to best measure the quality of an explanation? Furthermore, how strongly should we infer the best-explaining hypothesis or the best-explaining hypotheses? A number of competing measures of explanation quality have thus been put forward, see Section 2.2. We contribute to this ongoing work of determining the best, or most appropriate in a given situation, measure of explanation quality. Qualities of explanations have also been considered in formal ("Dung-style") argumentation models [17] and black-box models [18].

Standard interpretations of IBE currently tell us to infer the hypothesis that best explains the data, they do however not specify how strongly we should infer or believe the hypothesis that best explains the data. In this paper, a hypothesis of interest is a specific set of parameters in an Agent-Based Model (ABM), and inferring the parameters that best explain asset prices seems unwarranted. In order to avoid overfitting, it appears more prudent instead to infer various hypotheses to the degrees to which they explain the

* Corresponding author at: Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich, Germany.
E-mail address: juergen_landes@yahoo.de (J. Landes).

https://doi.org/10.1016/j.ijar.2025.109388
Received 15 October 2024; Received in revised form 15 January 2025; Accepted 17 February 2025

Available online 20 February 2025

available data. We show hence how IBE and the Bayesian framework can be intertwined to obtain probabilities for parameter values and subsequently evaluate these different methods for determining posterior probabilities via ABM simulations.[1]

ABMs are commonly used computational tools in several scientific disciplines ranging from biology [20] and epidemiology [21] to economics [22] and the social sciences [23]. They are simulation-based models that allow the study of collective phenomena emerging from the interactions of autonomous agents. ABMs typically require the specification of a number of parameters, some of which do not have equivalents in the real world. The development of general methods to find these parameters using real-world data, a process commonly referred to as "calibration" or "estimation", is a current challenge and an active area of research, especially for economic ABMs [24–28]. Many of these methods require the definition of meaningful priors, which however might sometimes be arbitrary and thus raise concerns about the suitability of posterior probabilities. In this context, the incorporation of further information into the posterior via explanatory relationships may be seen as carrying some promise. We here investigate this promise in the context of ABMs, which are closer to real-world applications than the previously used synthetic evaluation environments.

Our main contributions are i) an improved evaluation of the measures of explanation quality in [7] and ii) a proposal for a model of gradual interaction between Bayesian and IBE probabilities improving on [29,30].

The rest of the paper is organised as follows: We next provide background and describe our methodology (§ 2), discuss its implementation (§ 3) and then report and interpret the results of our simulations (§ 4). § 5 discusses our results and offers some conclusions.

2. Background and methodology

In this section, we outline the theoretical and methodological foundations of this study. First, the Background subsection introduces IBE, agent-based models and agent-based modelling. Next, the section Measures of Explanation Quality discusses various methods for evaluating the explanatory power of hypotheses, which play a central role in determining their plausibility. The subsection titled Evidence addresses how data, specifically time series, are processed and discretised to avoid computational difficulties. In the section Specification of Prior Probabilities, we explain how prior beliefs are assigned to the model parameters, utilising simulations to compute likelihoods. Finally, the section Posterior Probabilities explores different methodologies for updating beliefs, comparing the Bayesian and IBE approaches for determining posterior distributions over hypotheses.

2.1. Background

Here, we address two fundamental concepts for the building of this article: IBE and ABM. IBE is a reasoning approach where one infers the most likely hypothesis based on how well it explains the available data. It has been widely discussed across disciplines, particularly in relation to and in comparison with Bayesian methods. ABMs, on the other hand, are a computational approach used to simulate complex economic systems by modelling the interactions of diverse agents, offering valuable insights into emergent behaviours such as market dynamics, business cycles, and financial instability. Since these two concepts form the foundation of this article, we will further elaborate on both in the following sections.

2.1.1. Inference to the best explanation

People reason non-deductively from observations in a number of ways. Explanations play central parts in our cognitive processes, e.g., how we change our minds [31]. That a hypothesis or theory explains what is before our eyes tends to raise, in some shape or form, our confidence in the hypothesis/theory.

What one – or an artificial agent – *ought* to infer from the available information, evidence and/or data has been debated for many decades in a variety of fields including economics, computer science, philosophy and cognitive science. One important theme is that ideal (normative) rationality must resemble the most sophisticated reasoners (human beings) to some degree. Since we pay much attention to explanations and cherish their roles in learning, teaching and reasoning, it is not surprising that explanations have also been appealed to in normative theories of rationality. The theories we are concerned with here are known under the name *Inference to the Best Explanation* which is covered by the umbrella term *abduction*.

The basic idea is simple enough, one ought to believe what best explains the data. The typical story goes as follows. The cheese left on the kitchen table yesterday evening is gone in the morning and there is a small hole in the wall. These observations are best explained by a mouse finding a late-night snack. Your goldfish devouring the cheese is a much worse explanation, since the fish lives in the garden pond and is rather ill-equipped to make a hole in the wall. So, one ought to infer that the mouse and not the goldfish ate the cheese. Summarising, abduction is "*a mode of inference that makes explanatory considerations relevant to what we are licensed to believe*" [14, Section 1.2.1].

A number of questions arise immediately: How do we know what a good explanation is, what even *is* an explanation? What if a hypothesis explains perfectly but is otherwise ridiculous? What if two competing hypothesis explain equally well? What if all hypotheses explain poorly? How strongly should one believe a well-explaining hypothesis compared to a hypothesis which explains less well? What is the relationship to Bayesian rationality?

We refer the reader for answers to the first question to the debate on the *nature* of explanations to [32–35], which has become of interest in explainable AI [36–38]. The second type of question can be raised from a Bayesian point of view insisting on the importance of the prior probability of hypotheses; see [3, Section 2] for further discussions. We will return to the other four types of questions in

[1] Our comparison of IBE posteriors with the Bayesian approach is thus data-driven rather than based on a priori arguments, see [19] for the latter approach.

the following. A word of warning to those unfamiliar with IBE and abduction: proponents of this approach acknowledge violations of dynamic Dutch book arguments but deny that this is fatal and instead highlight attractive properties [3,29]. The relationship between IBE and Bayesian rationality has also been investigated in [39].

IBE can, in principle, be applied to all situations of uncertain reasoning. We however think that there is little to be gained in situations in which the data clearly indicate the actual state of the world, e.g., via uncertainty resolution [40], Bayesians refer to this situation as a "washed out prior" [41], statisticians think of this as successful "calibration" [42] and successful applications of Big Data more recently. If however there is much less information and uncertainty is ripe, then there seems to be a larger potential gain by exploiting explanatory relationships for formalising beliefs. In the following, we are considering the problem of assigning prior beliefs to parameters in an agent-based model. Since we are highly unsure about these parameters, we adopt a uniform prior. We hence believe that an application of IBE might help us to find better probabilities.

2.1.2. *Economic agent-based models*

Agent-Based Modeling (ABM) is a computational approach that has gained significant traction in the analysis of complex economic systems. Unlike traditional economic models that often rely on aggregate variables and assume representative agents with homogeneous behaviours, ABM focuses on the interactions of heterogeneous, autonomous agents. These agents operate based on individual rules and adapt their behaviours in response to interactions with other agents and changes in their environment [43]. At its core, ABM involves the creation of virtual environments where individual agents - such as consumers, firms, and policymakers - interact according to predefined rules. These interactions can lead to the emergence of complex macroeconomic phenomena that are not easily predictable from the behaviour of individual agents alone. The strength of ABM lies in its ability to model bottom-up dynamics, capturing how micro-level behaviours aggregate to produce macro-level outcomes [44]. Agents in ABM are typically endowed with characteristics such as memory, learning capabilities, and decision-making processes that can evolve over time. This allows ABM to incorporate concepts like bounded rationality, where agents make decisions based on limited information and cognitive resources, reflecting more realistic economic behaviours compared to the fully rational agents often assumed in traditional models [45,46]; for a game-theoretic discussion see also [47, Section 2].

The application of ABM in economics has provided new insights into various economic phenomena that are difficult to analyse using conventional models. One of the primary areas where ABM has been influential is in the study of market dynamics and financial systems. By simulating the interactions of individual market participants, ABM can capture the emergence of market prices, trading volumes, and volatility patterns that result from the collective behaviour of agents [48].

ABM has also been instrumental in exploring the mechanisms behind business cycles and economic fluctuations. Traditional economic theories often rely on equilibrium conditions to explain cycles, whereas ABM allows for the examination of out-of-equilibrium dynamics and the role of agent heterogeneity in generating economic instability [43]. For instance, the interaction of firms with varying strategies and adaptive behaviours can lead to endogenous cycles of growth and recession, providing a more nuanced understanding of economic volatility.

In the financial sector, ABM has been applied to model the behaviour of investors, financial institutions, and markets. These models can simulate scenarios such as asset price fluctuations, liquidity crises, and the effects of regulatory policies on financial stability. By incorporating features such as adaptive learning and bounded rationality, ABM provides a more realistic representation of market participants' decision-making processes compared to traditional models that assume perfect rationality [49].

ABM has been particularly useful in understanding systemic risk and the propagation of financial shocks. By modelling the interconnectedness of financial institutions and the feedback loops between them, ABM can identify potential sources of instability and assess the effectiveness of regulatory interventions designed to mitigate systemic risk [50]. This capability is crucial for designing policies that enhance the resilience of financial systems against crises.

Additionally, ABM has been employed to study market microstructure and the impact of high-frequency trading. By simulating the interactions between different types of traders, including algorithmic and human traders, ABM can analyse how trading strategies influence market liquidity, price discovery, and volatility [51]. The pioneering work by Brock and Hommes [52] on asset pricing will be further described in the following sections.

Although ABM can provide valuable insights for regulators and market participants supporting fair and efficient market exchanges, ABM also presents several challenges. The calibration and validation of agent-based models require detailed data on individual behaviours and interactions, which can be difficult to obtain. Moreover, the computational intensity of simulating large-scale ABMs can be a limitation, although advances in computational power and specialised software tools have mitigated this issue to some extent. Another challenge is ensuring the robustness and reliability of computations. Due to the inherent complexity and stochastic nature of these models, it is essential to perform extensive sensitivity analyses and to verify that the results are not artefacts of specific model configurations or parameter choices [43]. Transparency in model specification and rigorous methodological standards are crucial for the credibility and reproducibility of ABM research. The present work seeks to advance knowledge in this direction. As will be detailed further, we aim to explore inference to the best explanation methods for ABMs, aligning with the research strand of data-driven ABMs (see, for instance, [53]).

2.2. *Measures of explanation quality*

Our main aim is to determine reasonable probabilities over sets of parameters specifying ABMs. Parameters are real-valued and thus there are potentially an infinite number of combinations of them. To circumvent this issue, we choose to discretise the parameter space into a finite number of bins. We hence aim to assign probabilities to the elements of a set of hypotheses $\mathbb{H} = \{H_1, \dots, H_n\}$, which we take to be a *partition*, meaning that the H_i are mutually exclusive and exhaustive. We refer to [54,55] for measures of explanation quality for hypotheses that do not form a partition and to [56,57] for explanations from conjunctions of hypotheses.

Measures of Explanation Quality aim to formalise the intuitive idea that some hypotheses do a better job of explaining what is before our eyes. What exactly an explanation is, shall not concern us here. Instead, we consider the measures in [7], which assign every hypothesis $H \in \mathbb{H}$ an explanatory quality score given the available evidence, E, see [58,59] for more background.

These measures are $\mathcal{E}_{SS}$ (Schupbach & Sprenger, [60]), $\mathcal{E}_{CT}$ (Crupi & Tentori, [10]), $\mathcal{E}_{GM}$ (Good & McGrew, [61,62]), $\mathcal{E}_{OCM}$ (Overlap Coherence Measure, [5,6]), $\mathcal{E}_{PCM}$ (Product Coherence Measure, [63]), $\mathcal{E}_{LR}$ (Likelihood Ratio), $\mathcal{E}_{MPE}$ (Most Probable Explanation) and $\mathcal{E}_{ML}$ (Maximum Likelihood)[2] and are defined as follows:

$$\mathcal{E}_{SS}(H,E) = \frac{P(H|E) - P(H|\neg E)}{P(H|E) + P(H|\neg E)}$$

$$\mathcal{E}_{CT}(H,E) = \begin{cases} \dfrac{P(E|H) - P(E)}{1 - P(E)}, & \text{if } P(E|H) \geq P(E) \\ \dfrac{P(E|H) - P(E)}{P(E)}, & \text{if } P(E|H) < P(E) \end{cases}$$

$$\mathcal{E}_{GM}(H,E) = \ln\left(\frac{P(E|H)}{P(E)}\right)$$

$$\mathcal{E}_{OCM}(H,E) = \frac{P(H \wedge E)}{P(H \vee E)}$$

$$\mathcal{E}_{PCM}(H,E) = P(E|H) \cdot P(H|E)$$

$$\mathcal{E}_{LR}(H,E) = \frac{P(E|H)}{P(E|\neg H)}$$

$$\mathcal{E}_{MPE}(H,E) = P(H|E)$$

$$\mathcal{E}_{ML}(H,E) = P(E|H)\ .$$

These measures are, as their names suggest, meant to capture how well a hypothesis explains data. Note that these measures of explanation quality are not additive: $\mathcal{E}(H_1 \vee H_2, E)$ is in general not equal to $\mathcal{E}(H_1, E) + \mathcal{E}(H_2, E)$. There is a tradition of arguing about the (relative) merits of these measures on a priori grounds, see for example [7,63,64].

Since $\mathcal{E}_{SS}, \mathcal{E}_{CT}$ and $\mathcal{E}_{GM}(H,E)$ can take negative values and we want to combine measures of explanation quality with Bayesian approaches, they do not lend themselves directly to our current purposes. There are, of course, ways in which these measures can be changed to produce positive measures of explanation quality such as adding a suitably large positive constant or via exponentiation. Such normalisation procedures invalidate the axiomatic characterisations [5,6,10,60–63] motivating these measures. Furthermore, since different normalisation procedures result in markedly different measures, it is not clear at all which normalisation to apply. In the case of $\mathcal{E}_{GM}$ however, there is a canonical normalisation procedure by dropping the logarithm, which is a normalisation function mapping the positive real numbers to the set of all real numbers. Since there is no canonical way to normalise $\mathcal{E}_{SS}, \mathcal{E}_{CT}$ we will not consider these measures in the following. We will come back to the normalised $\mathcal{E}_{GM}$ measure, denoted by $\mathcal{E}_{GMN}$, in Section 3.1.3.

2.3. Evidence

The ABMs we use in our simulations depend on multiple parameters. For every fixed set of parameters they non-deterministically output a time series of asset prices. For every fixed set of parameters, the conditional probability of obtaining a particular time series is (very close to) zero, if the time series spans more than a few discrete time steps. Conditional probabilities that are (too close to) zero create computational intractabilities (rounding errors swamping data and divisions by zero).

In order to circumvent this computational problem, we do not take these time series as our observations but rather the discretised first and second moments of the time series. Provided we have established bounds on their mean and the variance, the number of values the observations, E, may take is finite. Let us denote this set by $\mathbb{E}$.

2.4. Specification of prior probabilities

Prior probabilities have to be assigned to the space spanned by $\mathbb{H}$ and $\mathbb{E}$, $\mathbb{H} \times \mathbb{E}$. We do this in the usual way by i) assigning prior probabilities to the parameter space, $P(H)$ for all $H \in \mathbb{H}$ and ii) setting conditional probabilities (likelihoods), $P(E|H)$ for all $E \in \mathbb{E}$ and all $H \in \mathbb{H}$.[3]

Due to the complexity of the ABMs, the conditional probabilities $P(E|H)$ (likelihoods) cannot be calculated analytically. Rather, we can compute them numerically as natural frequencies observed in simulations. Recall that every fixed hypothesis H is determined

[2] Glass adopts $\mathcal{E}_{MPE}(H,E)$ as a baseline rather than a comprehensive IBE measure (see [7, footnote 2] and [63]). Moreover, his focus on an ordinal ranking means that he does not distinguish between the first three measures, as they yield identical rankings (see again [7]). However, since we are interested in concrete values, we cannot conflate these measures and treat them individually.

[3] In [7], Glass also considered performances of posteriors depending on prior probabilities. Since we did not find intuitively reasonable different priors in our modelling approach, we have not considered different prior probabilities.

by a set of parameter values. For these fixed parameter values we can run the models and observe the distribution of the statistical properties (here, first and second moments).

2.5. Posterior probabilities

Recall that we want to compare different methodologies for determining sensible posterior probabilities over the hypothesis space $\mathbb{H}$. A fully Bayesian analysis is often – as it is here – possible. However, since we lack a sensible way for setting priors (in our implementation we resort to adopting the uniform prior on $\mathbb{H}$), the degree to which the ABMs explain statistical properties of the observed data may provide extra information [16]. We shall thus compare the Bayesian approach with IBE methods.

Douven and Wenmackers suggested in [29,30] the following formula for obtaining a posterior probability over $\mathbb{H}$ taking the explanatory relationships between hypotheses and evidence, E, into account

$$P_f(H) = \frac{P(H) \cdot P(E|H) + f(H,E)}{\sum_{H_i \in \mathbb{H}} P(H_i) \cdot P(E|H_i) + f(H_i,E)} \ , \tag{1}$$

where f is a function rewarding only the best explanation

$$f(H) = \begin{cases} c, & \text{if } H \text{ best explains } E \\ 0, & \text{otherwise} \ . \end{cases}$$

If $c = 0$, then the function f is equal to zero for all inputs and the updating procedure is the usual Bayesian posterior on $\mathbb{H}$, which we denote by P_0.

Recall that measures of explanation quality are not additive, in general. The functions defined in Equation (1) are hence, in general, also not additive. In order to define a posterior *probability* over $\mathbb{H}$ we *define* the posterior probability, P_f, of a disjunction of hypotheses, as the sum of the posteriors of hypotheses: $P_f(\bigvee_{q \in Q \subseteq \{1,\ldots,n\}} H_q) := \sum_{q \in Q \subseteq \{1,\ldots,n\}} P_f(H_q)$. This convention shall also apply to the functions defined in Equations (2) and (3).

We see two major issues and one minor problem with the suggestion by Douven and Wenmackers. Firstly, the value of the constant c is left unspecified. Although, possible values of c are explored in a number of simulations in [30], it remains unclear which value of c is most appropriate in their simulations on a fixed toy domain of coin tossing. This means that the value of c for our application is not clear at all. Secondly, f rewards only the best explanation. The degree to which other hypotheses explain the data is irrelevant. Such a strong condition seems unwarranted for our ABMs, because differences in explanatory power can be minimal. Finally, it is not obvious what to do when several hypotheses explain the evidence equally well. Should they all get the reward c, or should the reward be distributed equally among them?

Bayes + IBE In order to side-step these complications we modified the proposed formula in [29,30] by an alternative reward function, which provides gradual rewards and does not require the specification of a rather arbitrary parameter. Let us thus fix a measure of explanation, $\mathcal{E} \notin \{\mathcal{E}_{SS}, \mathcal{E}_{CT}\}$, and consider the degree to which hypotheses explain the evidence E

$$P_{\text{Bayes+IBE},\mathcal{E}}(H) := \frac{P(H) \cdot P(E|H) + \mathcal{E}(H,E)}{\sum_{H_i \in \mathbb{H}} P(H_i) \cdot P(E|H_i) + \mathcal{E}(H_i,E)} \ . \tag{2}$$

The above convention on disjunctions extends these functions to disjunctions of hypotheses.

Having excluded problematic measures of explanation quality (negative values, see § 2.2), we note that $P_{\text{Bayes+IBE},\mathcal{E}}$ is a probability distribution over $\mathbb{H}$.

IBE A proponent of pure IBE might not be satisfied with the mixing of Bayesian posterior probabilities with explanation qualities in [29,30] and (2). To answer this possible criticism, we shall also consider a purely IBE-based way of gradually assessing hypotheses:

$$P_{\text{IBE},\mathcal{E}}(H) := \frac{\mathcal{E}(H,E)}{\sum_{H_i \in \mathbb{H}} \mathcal{E}(H_i,E)} \ . \tag{3}$$

The above convention on disjunctions extends these functions to disjunctions of hypotheses.

Having excluded problematic measures of explanation quality (see § 2.2), we note that $P_{\text{IBE},\mathcal{E}}$ is a probability distribution over $\mathbb{H}$.

Since the Most Probable Explanation posteriors, $P_{\text{IBE},\mathcal{E}_{MPE}}(H)$ and $P_{\text{Bayes+IBE},\mathcal{E}_{MPE}}(H)$ are both equal to the Bayesian posterior of H,

$$P_{\text{IBE},\mathcal{E}_{MPE}}(H) = \frac{P(H|E)}{\sum_{H_i \in \mathbb{H}} P(H_i|E)} = \frac{P(H|E)}{1} = P(H|E)$$

$$P_{\text{Bayes+IBE},\mathcal{E}_{MPE}}(H) = \frac{P(H) \cdot P(E|H) + P(H|E)}{\sum_{H_i \in \mathbb{H}} P(H_i) \cdot P(E|H_i) + P(H_i|E)}$$

$$= \frac{P(E \wedge H) \cdot (1 + \frac{1}{P(E)})}{\sum_{H_i \in \mathbb{H}} P(E \wedge H_i) + \frac{P(E \wedge H_i)}{P(E)}}$$

$$= \frac{P(E \wedge H) \cdot (1 + \frac{1}{P(E)})}{P(E) \cdot (1 + \frac{1}{P(E)})} = \frac{P(E \wedge H)}{P(E)} = P(H|E) \ ,$$

we exclude them from further considerations. Note that these posteriors are equal for all prior probability function defined on $\mathbb{H}$.

We have thus defined 10 different ways ($P_{\text{Bayes+IBE},\mathcal{E}}$ and $P_{\text{IBE},\mathcal{E}}$ for 5 different measures $\mathcal{E}$: $\mathcal{E}_{GMN}, \mathcal{E}_{OCM}, \mathcal{E}_{PCM}, \mathcal{E}_{LR}, \mathcal{E}_{ML}$) of setting posterior probabilities by applying IBE. We go on to compare these 10 methods with the traditional Bayesian way of obtaining the posterior by conditionalisation, P_0.

The 5 explanation measures (the $\mathcal{E}$) of current interest can be computed from the prior (the $P(H)$ for $H \in \mathbb{H}$) and likelihoods (the $P(E|H)$ for $H \in \mathbb{H}$) as follows. The prior is flat. So, $P(H) = P(H_i)$ for all $H, H_i \in \mathbb{H}$. Furthermore,

$$\begin{aligned}
\mathcal{E}_{LR}(H,E) &= \frac{P(E|H)}{P(E|\neg H)} = \frac{P(E|H) \cdot P(\neg H)}{P(E \wedge \neg H)} \\
&= \frac{P(E|H) \cdot [1 - P(H)]}{\sum_{H_i \in \mathbb{H}\setminus\{H\}} P(E|H_i) \cdot P(H_i)} \\
\mathcal{E}_{OCM}(H,E) &= \frac{P(H \wedge E)}{P(H \vee E)} = \frac{P(H \wedge E)}{P(H) + P(E \wedge \neg H)} \\
&= \frac{P(H) \cdot P(E|H)}{P(H) + \sum_{H_i \in \mathbb{H}\setminus\{H\}} P(H_i \wedge E)} \\
&= \frac{P(H) \cdot P(E|H)}{P(H) + \sum_{H_i \in \mathbb{H}\setminus\{H\}} P(E|H_i) \cdot P(H_i)} \\
\mathcal{E}_{PCM}(H,E) &= P(E|H) \cdot P(H|E) = \frac{P(E|H) \cdot P(H \wedge E)}{P(E)} = \frac{P(E|H)^2 \cdot P(H)}{\sum_{H_i \in \mathbb{H}} P(H_i \wedge E)} \\
&= \frac{P(E|H)^2 \cdot P(H)}{\sum_{H_i \in \mathbb{H}} P(E|H_i) \cdot P(H_i)} \ .
\end{aligned}$$

We will show in Section 3.1.3 that the other two explanation measures lead to posterior probabilities that are equal to the Bayesian posterior and thus trivially computable from the prior.

Before moving on, we briefly discuss properties of the approach suggested here. Following the spirit of the IBE approach, the posteriors calculated here are based on *all the available evidence*. Splitting the body of available evidence into multiple parts and updating piecemeal is thus undesirable, it will also produce, in general, results that depend on the way the evidence is split. Our approach thus differs in this respect from traditional Bayesian updating via conditionalisation but resembles an objective Bayesian approach via maximising entropy given all the available evidence [65,66].

We use a uniform prior over the space of hypothesis here, which reflects our lack of background knowledge. In other circumstances, background knowledge is available and can be plugged into our approach via a non-uniform prior.

3. Implementation

We next describe the implementation of our testing environment, beginning with the setup based on the ABM developed by Brock and Hommes for asset pricing [52]. We explain the model's dynamics and parameter selection process based on prior validation literature. We then describe the evidence production process, which involves generating simulated time series and calculating Bayesian and inference to the best explanation (IBE) posteriors. Finally, we introduce two performance measures for assessing parameter approximation and forecast accuracy. Table A.1 provides a pseudo-code description of our implementation.

3.1. Set up

This section provides an overview of the key components in our setup. We employ the Brock and Hommes ABM to simulate asset pricing, focusing on parameter adjustments for realistic outcomes. We generate time series data, use first and second moments as evidence, and discretise the evidence space with a hexagonal tessellation. We then compute posterior probabilities for various hypotheses using Bayesian, IBE, and hybrid approaches.

3.1.1. The agent-based model and its parameters

For our testing purposes, we employed the closed-form ABM model developed by Brock and Hommes for asset pricing [52]. This model is well known in the literature as an early example of a class of ABMs that simulate the trading of assets on an artificial stock market by modelling the interactions of heterogeneous traders following various trading strategies.

Every strategy, s, has an associated trend-following component, g_s, and bias, b_s, both of which are real-valued parameters that are of particular interest in our investigation. The number of strategies, S, can vary depending on the application. In line with previous work [25,52,67,68], we selected $S = 4$. Thus, there are four trading strategies (BH4) in this model, each specified by two real parameters, g_i and b_i for $i \in \{1, 2, 3, 4\}$, where again g_i represents trend-following and b_i represents bias. The model dynamics

Table 1
Parameters specification for Setting 1 and Setting 2 as in [25,28]. Note that in our experiments g_2 and b_2 are the two free parameters.

Parameters	Setting 1	Setting 2
(g_1, b_1)	$(0,0)$	$(0,0)$
(g_2, b_2)	$(0.9, 0.2)$	$(-0.7, -0.4)$
(g_3, b_3)	$(0.9, -0.2)$	$(0.5, 0.3)$
(g_4, b_4)	$(1.01, 0)$	$(1.01, 0)$

can be expressed via the following system of coupled equations, where the main component $x_{t+1} = p_{t+1} - p^*_{t+1}$, i.e., the deviation of the asset price p_t from its benchmark fundamental p^*_t is:

$$x_{t+1} = \frac{1}{R}\left[\sum_{s=1}^{S} n_{s,t+1}\left(g_s x_t + b_s\right) + \epsilon_{t+1}\right], \; \epsilon_t \sim \mathcal{N}(0,\sigma^2), \tag{4}$$

$$n_{s,t+1} = \frac{\exp\left[\beta\left(x_t - Rx_{t-1}\right)\left(g_s x_{t-2} + b_s - Rx_{t-1}\right)\right]}{\sum_{s'=1}^{S}\exp\left[\beta\left(x_t - Rx_{t-1}\right)\left(g_{s'} x_{t-2} + b_{s'} - Rx_{t-1}\right)\right]} \;. \tag{5}$$

The parameters R, β and σ represent respectively: (a) the gross return, (b) the intensity of choice measuring how fast agents switch between different prediction strategies, and (c) the common, constant conditional variance (namely σ^2) that all traders are assumed to have. We follow [25] setting their values: $(R, \beta, \sigma) = (1.01, 10, 0.04)$.

We adopt the simulation environment outlined in [28, § 4.2, 4.2.1 and 4.2.2], which defines two settings by specifying parameter sets as described in Table 1. The two settings specified in the table were chosen to conform to previous literature. Our implementation employs the Python package `Black-it` [69].

We here took g_2 and b_2 as free parameters each ranging over the unit interval following [25,28]. Exploratory simulations showed that most parameter values for g_2 and b_2 produced (average) time series that were in line with how asset prices tend to behave, however some (combinations of parameters) produced outliers. We hence constrained both parameters to the range $[0.1, 0.9]$, which produced the desired effects of removing the outliers and being more in line with observed asset prices, see Fig. 1.

Discretising the interval $[0.1, 0.9]$ into steps of 0.05, we obtained $17^2 = 289$ pairs of parameters, where each pair represents a trading strategy. For both settings (fixed parameters $b_1, g_1, b_3, g_3, b_4, g_4$) we thus considered 289 competing hypotheses.[4]

3.1.2. Evidence production

The following procedure was carried out for *Setting 1* and *Setting 2*.

We sampled from every hypothesis 10,000 times (a total of $10^4 \cdot 289 = 2{,}890{,}000$ time series) for 200 timesteps each. As explained in Section 2.3, we here considered the first and second moment of the time series as the evidence, see Fig. 2 for sample time series.

For every observed time series we computed the average asset price over the 200 timesteps. For every hypothesis, $H \in \mathbb{H}$, we then averaged the 10,000 averaged asset prices obtaining an average first moment. Furthermore, for every observed time series we computed its second moment. For every hypothesis, $H \in \mathbb{H}$, we then averaged the 10,000 second moments obtaining an average second moment.

These averaged first and second moments fell into the intervals $[-1, 19]$ and respectively $[0, 140]$, cf. Fig. 1. The space of average first and second order moments, $\mathbb{E}$, was then covered by a hexagonal tessellation of 566,705 tiles. We could thus compute likelihoods for all hypotheses $H \in \mathbb{H}$ and all potential observations, $E \in \mathbb{E}$, see Fig. 3 for an example.

In order to obtain a single observed event $E \in \mathbb{E}$ on which to update, we calculated, for a fixed hypothesis H, the outcome with maximal likelihood $\arg\sup_{E\in\mathbb{E}} P(E|H) =: E^*$, cf. Fig. 4. There we plot the likelihoods $P(E^*|H)$ for five different $H \in \mathbb{H}$ in Setting 2. So, E denotes a potential observation and E^* the actual observation.

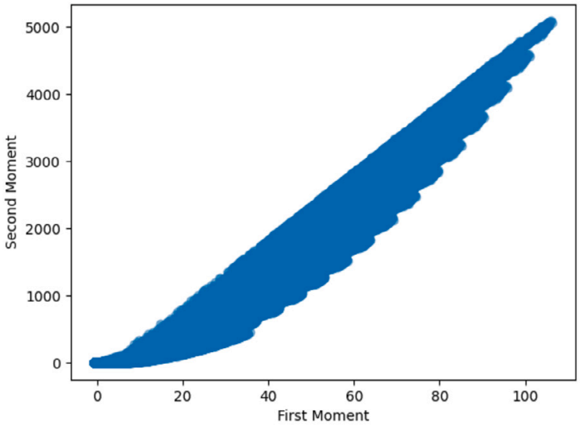
3.1.3. Updating

For *Setting 1* and *Setting 2* we chose these five pairs of parameters, g_2, b_2, representing the actual hypothesis, H^*: (0.25,0.25), (0.25,0.75), (0.5,0.5), (0.75,0.25) and (0.75,0.75). For each of these five pairs parameters, we used the outcomes of the simulations described in Section 3.1.2 as evidence. For these outcomes we also computed the likelihoods of the actual observations.

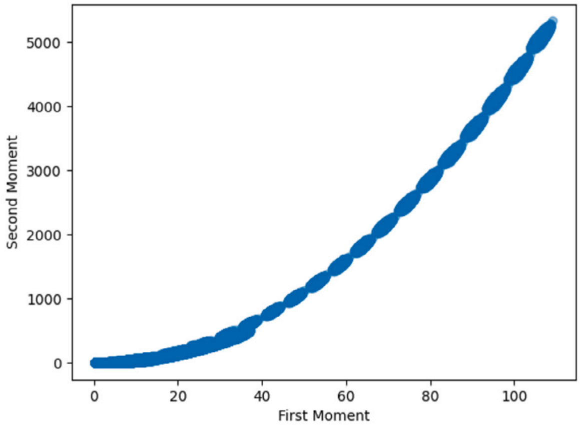
For *Setting 1* and *Setting 2*, for all $\mathcal{E} \in \{LR, OCM, PCM\}$ and all five pairs of parameters we computed the two types posterior probabilities $P_{\text{post}} : \mathbb{H} \to [0,1]$ described by (2) and (3)

$$P_{\text{IBE},\mathcal{E}}(H) := \frac{\mathcal{E}(H, E^*)}{\sum_{H_i\in\mathbb{H}} \mathcal{E}(H_i, E^*)}$$

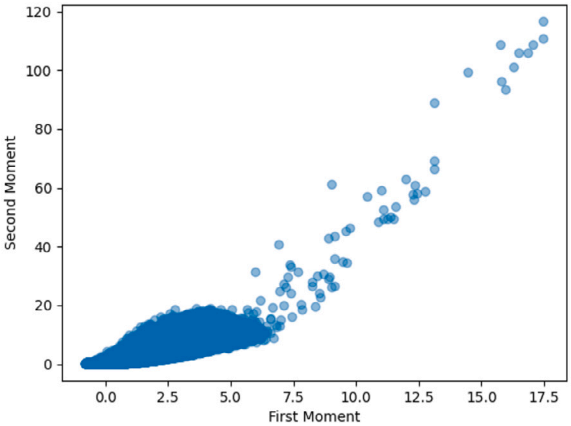
[4] This set-up entails that we only compare Brock and Hommes models with each other. We thus exclude from the beginning the possibility of a "true" state of the world, which is not a Brock and Hommes model.



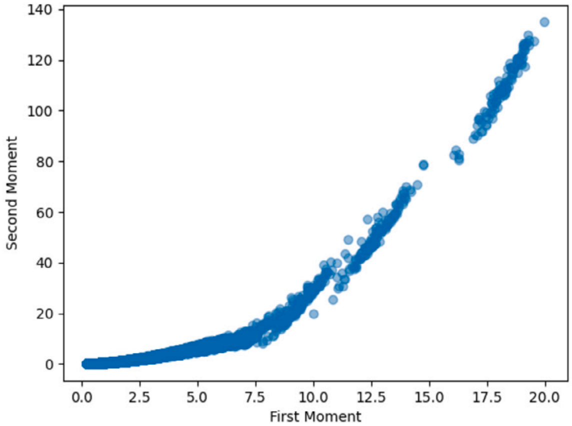
(a) Setting 1: $(g_2, b_2) \in [0, 1]$



(b) Setting 2: $(g_2, b_2) \in [0, 1]$

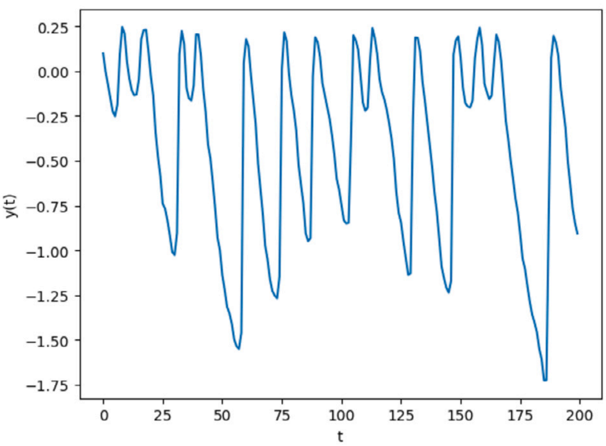


(c) Setting 1: $(g_2, b_2) \in [0.1, 0.9]$

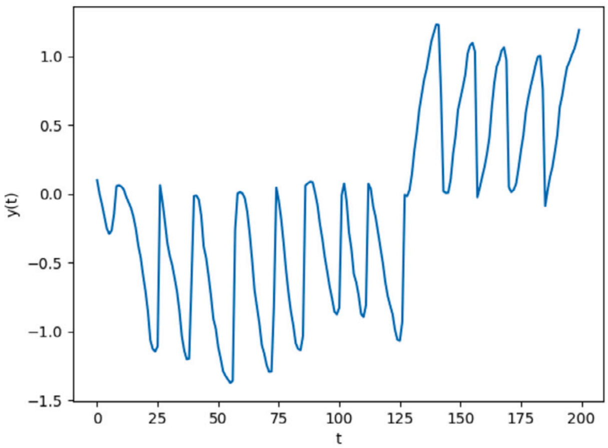


(d) Setting 2: $(g_2, b_2) \in [0.1, 0.9]$

Fig. 1. Scatter plot of first (X-axis) and second (Y-axis) moments of time series in the Brock and Hommes model. Every blue dot represents an average of 10,000 time series for 200 timesteps each, for a different pair of parameters (g_2, b_2). Top row: $(g_2, b_2) \in [0,1]^2$, lower row: $(g_2, b_2) \in [0.1, 0.9]^2$, left: Setting 1, right: Setting 2. The plots in the lower row show how restricting the pair (g_2, b_2) to $[0.1, 0.9]^2$ leads to more reasonable asset prices: a five-fold decrease in maximal average first moments and a 35-fold decrease in maximal average second moments. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

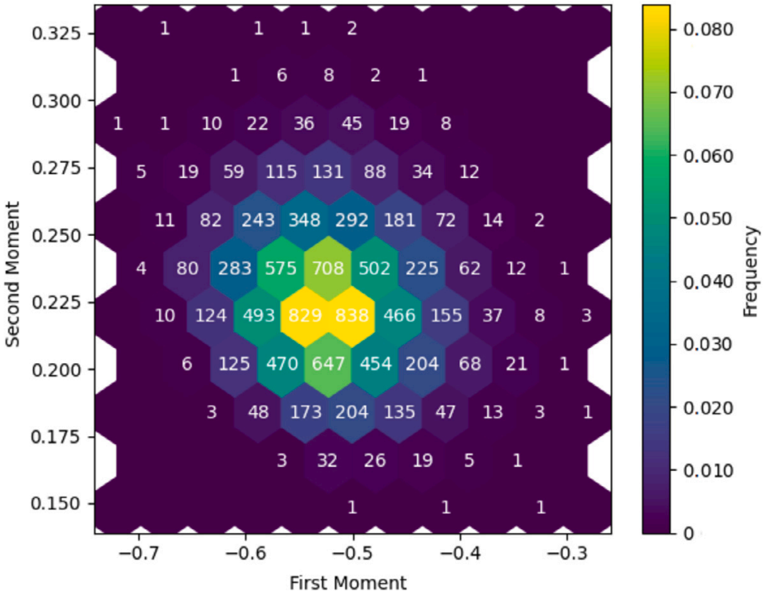


(a) $(g_2, b_2) = (0.1, 0.2)$

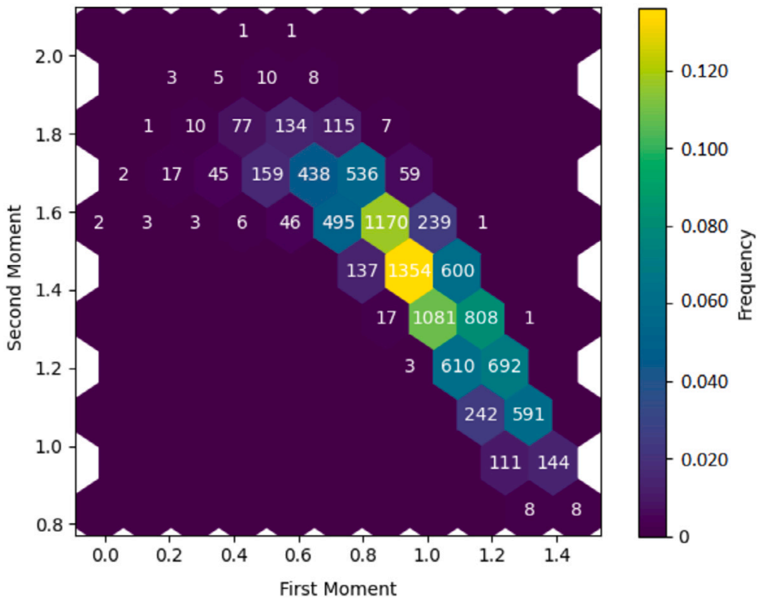


(b) $(g_2, b_2) = (0.9, 0.2)$

Fig. 2. Two examples of simulated target time-series originating from the Brock and Hommes model in Setting 1, where $y(t) = x_{t+1}$ for $t \geq 1$, see Equation (4).



(a) Setting 1



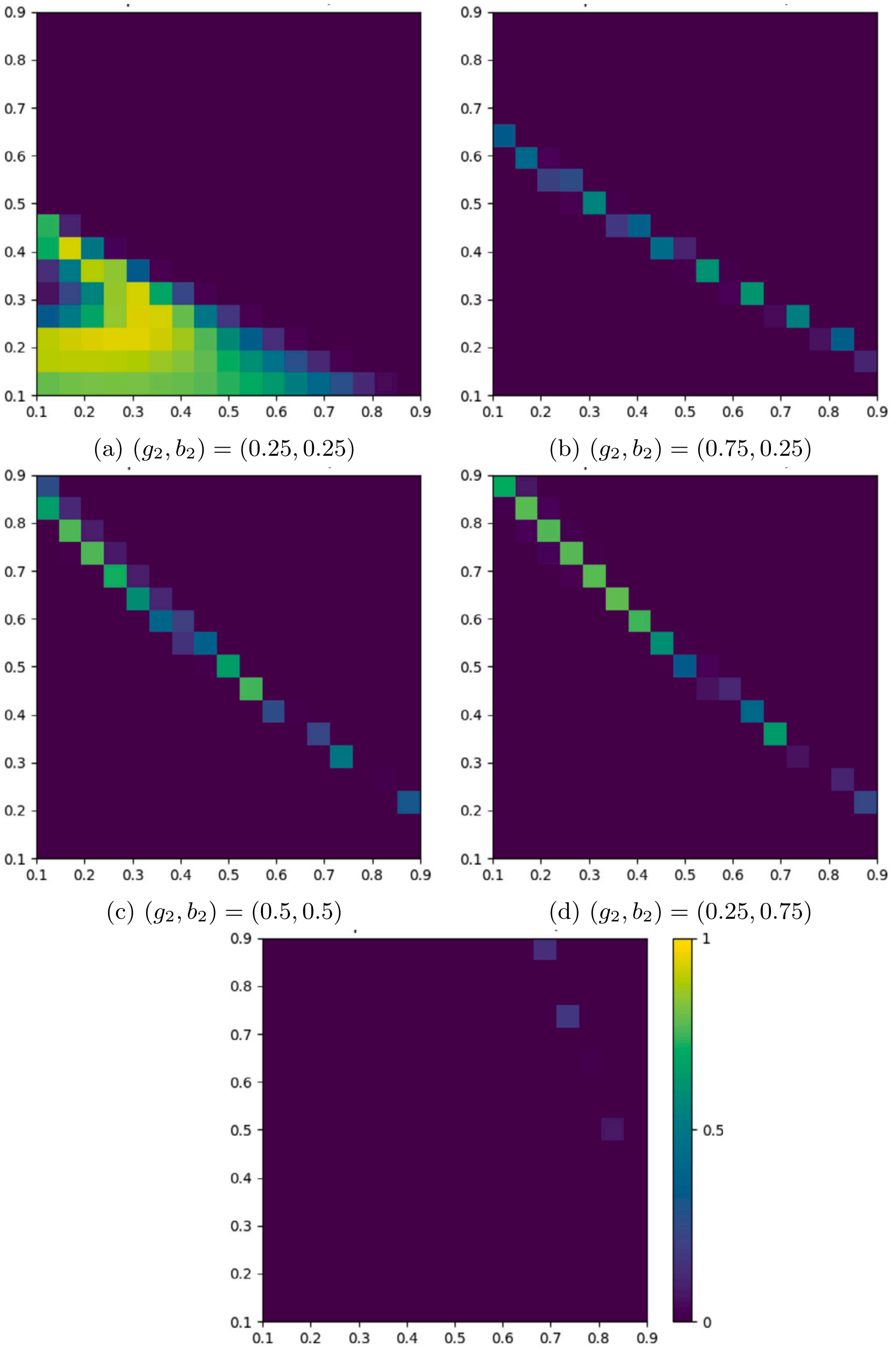
(b) Setting 2

Fig. 3. Two examples of sampling first (x-axis) and second (y-axis) moments in the tessellation. a) Setting 1 with $(g_2, b_2) = (0.8, 0.15)$, and b) Setting 2 (lower panel) with $(g_2, b_2) = (0.7, 0.75)$. The numbers indicate how many of the 10,000 observed time series have a first and second moment in a tile, tiles without a time series are not labelled.

$$P_{\text{Bayes+IBE},\mathcal{E}}(H) := \frac{P(H) \cdot P(E^*|H) + \mathcal{E}(H, E^*)}{\sum_{H_i \in \mathbb{H}} P(H_i) \cdot P(E^*|H_i) + \mathcal{E}(H_i, E^*)} \ .$$

We also computed the Bayesian posterior P_0 for all these cases.

Since the prior is flat, the maximum likelihood posteriors and the two normalised Good & McGrew posteriors all agree with the Bayesian posterior. Unlike the Most Probable Explanation posteriors, these posteriors are, in general, different, from the Bayesian posterior. These measures might be of interest in cases with a non-uniform prior.



(a) $(g_2, b_2) = (0.25, 0.25)$ (b) $(g_2, b_2) = (0.75, 0.25)$

(c) $(g_2, b_2) = (0.5, 0.5)$ (d) $(g_2, b_2) = (0.25, 0.75)$

(e) $(g_2, b_2) = (0.75, 0.75)$

Fig. 4. Likelihoods of the observed outcome with maximal likelihood: for a fixed H^* we computed the outcome with maximal likelihood $\arg\sup_{E\in\mathbb{E}} P(E|H) =: E^*$. We then plotted $P(E^*|H)$ for all $H \in \mathbb{H}$. All five plots are for Setting 2 for H^* given by g_2 (x-axes) and b_2 (y-axes) for varying (g_2, b_2).

$$
\begin{aligned}
P_{\mathrm{IBE},\mathcal{E}_{ML}}(H) &= \frac{P(E|H)}{\sum_{H_i\in\mathbb{H}} P(E|H_i)} \\
&= \frac{P(E\wedge H)}{P(H)} \cdot \frac{P(H)}{\sum_{H_i\in\mathbb{H}} P(E\wedge H_i)} \\
&= \frac{P(E\wedge H)}{P(E)} = P(H|E)
\end{aligned}
$$

$$
\begin{aligned}
P_{\text{Bayes+IBE},\mathcal{E}_{ML}}(H) &= \frac{P(H)\cdot P(E|H) + P(E|H)}{\sum_{H_i\in\mathbb{H}} P(H_i)\cdot P(E|H_i) + P(E|H_i)} \\
&= \frac{(P(H)+1)\cdot P(E|H)}{\sum_{H_i\in\mathbb{H}}(P(H_i)+1)\cdot P(E|H_i)} \\
&= \frac{P(E|H)}{\sum_{H_i\in\mathbb{H}} P(E|H_i)} \\
&= \frac{P(E\wedge H)\cdot P(H)}{P(H)\cdot\sum_{H_i\in\mathbb{H}} P(E\wedge H_i)} \\
&= \frac{P(E\wedge H)}{P(E)} = P(H|E) \\
P_{\text{IBE},\mathcal{E}_{GMN}}(H) &= \frac{\frac{P(E|H)}{P(E)}}{\sum_{H_i\in\mathbb{H}_i}\frac{P(E|H_i)}{P(E)}} = \frac{P(E|H)}{\sum_{H_i\in\mathbb{H}_i} P(E|H_i)} = P_{\text{IBE},\mathcal{E}_{ML}}(H) \\
&= P(H|E) \\
P_{\text{Bayes+IBE},\mathcal{E}_{GMN}}(H) &= \frac{P(H)\cdot P(E|H) + \frac{P(E|H)}{P(E)}}{\sum_{H_i\in\mathbb{H}}\left[P(H_i)\cdot P(E|H_i) + \frac{P(E|H_i)}{P(E)}\right]} \\
&= \frac{P(E|H)\left(P(H)+\frac{1}{P(E)}\right)}{\sum_{H_i\in\mathbb{H}} P(E|H_i)\left(P(H_i)+\frac{1}{P(E)}\right)} = P(H|E)\ .
\end{aligned}
$$

We hence exclude these measures from further considerations in this paper.

3.2. *Performance measures*

Next, we were interested in determining how well these posteriors performed. We therefore evaluated how well the posteriors managed to capture the five hypotheses specified by pairs of parameters g_2, b_2. We used two different measures.

3.2.1. *Measure 1: Parameters – distance on* $\mathbb{H}$

Firstly, we determined how well the posteriors approximate the true parameter values. This was achieved by computing the Euclidean distance, which is closely related to the Brier Score [70], between a posterior probability of hypotheses $P_{\text{post}} : \mathbb{H} \to [0,1]$ (as defined by (2) and (3) as well as P_0) and the probability, P^*, assigning one to the true parameter values and zero to all other parameter values. Recall that P_{post} depends on the actual observation and thus depends on the setting and the parameter pair g_2, b_2. Denoting the true hypothesis be H^* we have,

$$
\begin{aligned}
d_1(P_{\text{post}}, P^*) &:= \sqrt{\sum_{H\in\mathbb{H}}(P_{\text{post}}(H) - P^*(H))^2} \\
&= \sqrt{(1-P_{\text{post}}(H^*))^2 + \sum_{H\in\mathbb{H}\setminus\{H^*\}} P_{\text{post}}(H)^2}\ .
\end{aligned}
$$

3.2.2. *Measure 2: Forecasts – distance on* $\mathbb{E}$

The measure d_1 is invariant under all permutations of hypotheses σ that map H^* onto itself, $\sigma(H^*) = H^*$. This is undesirable in some sense, since the hypotheses, which are close to H^* in an Euclidean sense, tend to output similar time series, whereas hypotheses at the extremes of the hypothesis space tend to output time series that do not look at all like those of H^*.

Secondly, we also evaluated how well the posteriors approximate the *forecasts* of H^*. We thus determined the posterior probability of an $E \in \mathbb{E}$ as the product of the likelihood of E multiplied by the posterior of H,

$$
P_{\text{post}}(E) = \sum_{H\in\mathbb{H}} P_{\text{post}}(H)\cdot P(E|H)\ .
$$

The assumption of a uniform prior, $P(H) = \text{constant}$, ensures that the conditional probability $P(E|H)$ is well-defined.[5]

We then computed the Euclidean distance between these posterior probabilities and the likelihoods of E given H^* varying $E \in \mathbb{E}$. This distance is a measure of how well the predictions of the posteriors agree with the predictions of H^*. We have,

[5] Furthermore, we note that a normal Bayesian posterior probability of an elementary event is either zero or one. In our setting, the posterior probability is initially defined on $\mathbb{H}$ and serves as our new prior over $\mathbb{H}$, which can be used to define a “posterior” probability of $E \in \mathbb{E}$ in the way specified above.

$$d_2(P_{\text{post}}, P^*) := \sqrt{\sum_{E\in\mathbb{E}} (P^*(E|H^*) - P_{\text{post}}(E))^2}$$
$$= \sqrt{\sum_{E\in\mathbb{E}} \left(P(E|H^*) - \sum_{H\in\mathbb{H}} P_{\text{post}}(H) \cdot P(E|H) \right)^2} .$$

Recall from § 2.4 that prior probabilities P are defined as priors over the parameter space $\mathbb{H}$ and conditional probabilities (likelihoods) $P(E|H)$. § 3.1.2 explains that these likelihoods depend on E and $\mathbb{H}$ but not on P. Thus, $P^*(E|H^*) = P(E|H^*)$.

We refrain from defining a third measure, which measures distances of probabilities on $\mathbb{H} \times \mathbb{E}$. Such a measure would have to compare – and trade off – distances between parameter values and forecasts, since we did not wish to enter the treacherous grounds of such comparisons and trade-offs, cf. [71, p. 249].

3.2.3. Most likely hypotheses

One could argue that the spirit of IBE is to adopt a single hypothesis after receiving information. We have so far defined posterior probabilities and described how to measure their performances. Instead, one might want to (defeasibly!) infer the hypothesis with largest posterior probability. We thus computed the hypotheses with the largest posterior probability and determined on which hypothesis a posterior puts the most probability, see Tables 6 and 7.[6]

4. Results and interpretation

We now report and interpret the results of our simulations for the two settings, the two measures and the five parameter pairs (g_2, b_2), a total of 20 scenarios.

4.1. Results

Tables 2–5 report the performance of the studied posteriors for Settings 1-2 and Measure 1-2, respectively. In all Tables, best-performing posteriors are highlighted in yellow, posteriors performing better than the Bayesian posterior are highlighted in orange, and posteriors performing worse than the Bayesian posterior are coloured blue. Our noteworthy findings are as follows.

Performance Our most important finding is that there was no posterior which performs best in all scenarios. While there was no posterior that always performed best, the posterior P_{PCM} performed best in 15 out of 20 scenarios. The Bayesian posterior, often considered the gold-standard of rationality, performed best in 3 out of 20 scenarios. The PCM + IBE posterior as well as the LR posteriors performed best in one scenario. There was no scenario in which an OCM posterior performed best.

IBE vs. Bayes Overall, the IBE posteriors put forward in (3) performed better than the Bayesian posterior.

Mixing Bayes and IBE There was only one scenario in which a Bayes + IBE posterior performed better than all other posteriors, see Table 4. In all other cases, the Bayes + IBE Posteriors never outperformed the Bayesian *and* the respective IBE posterior. Instead, these mixed posteriors were closely aligned with their respective IBE posterior.

Bayes The Bayesian posterior did not outperform a single posterior in a scenario in Setting 2 according to Measure 1, see Table 4. Using Measure 1 instead, the Bayesian posterior only outperformed the PCM posteriors in one scenario.

ML, MPE and GMN The ML, MPE and GMN posteriors are all equal to the Bayesian posteriors and are hence not reported in the tables.

Measure The choice of a measure is relevant to the relative performance of posteriors. Tables 2 and 3 differ only in terms of the performance measure of posteriors as do Tables 4 and 5, however we did find differences in measured performances highlighted by different colourings in these tables. These findings highlight the importance of employing more than one evaluation methodology. Had we only considered Measure 1 and Setting 2, we would have concluded that the Bayesian approach is clearly worst and PCM is the best IBE approach, see Table 4. After carrying out multiple assessments, a different and more varied picture emerges.

Tables 6 and 7 show that in both settings and for all choices of H^* all posteriors assign the same hypothesis H the greatest probability. Only in the case of $(0.75, 0.75)$ do they assign the greatest probability to H^*. The pure PCM posterior always assigns the maximal probability to the a posterior most probable hypothesis, for every setting and every choice of H^* there exists a hypothesis $H_{\max} \in \mathbb{H}$ that has maximal posterior probability for every posterior probability, $H_{\max} = \arg\sup_{H\in\mathbb{H}} P_{\text{post}}(H)$, and $P_{\text{IBE},\mathcal{E}_{PCM}}(H_{\max}) > P_{\text{post}}(H_{\max})$ for all posteriors $P_{\text{post}} \neq P_{\text{IBE},\mathcal{E}_{PCM}}$.

4.2. Interpretation

Our results suggest that sweeping claims trumpeting the superiority of one approach over the other are not justified.

The relative performances of posteriors were sensitive to the setting (background parameters, $(g_1, b_1, g_3, b_3, g_4, b_4)$), the measure used to evaluate performance and the true state of the world to be learned (g_2, b_2). Furthermore, we used an Euclidean distance measure to assess performances of posterior probabilities. While Euclidean distance measure are often the measure of choice there is no means the only distance measure one could consider [65,72–75].

[6] This evaluation has been suggested to us by Sébastien Destercke.

Table 2
Measure 1: Setting 1 For $(g_2, b_2) = (0.25, 0.25), (0.25, 0.75), (0.75, 0.75)$ the posterior P_{PCM} outperforms all other posteriors (in terms of minimising d_1). For the other two parameter pairs $(g_2, b_2) = (0.5, 0.5), (0.75, 0.25)$ the Bayesian posterior, P_{ML} and P_{MPE} and their combinations outperform all other posteriors.
All posteriors performed worst on the parameters $(g_2, b_2) = (0.5, 0.5)$ and best on the parameters $(g_2, b_2) = (0.75, 0.75)$. While there is in general not much difference in performance between posteriors (for fixed (g_2, b_2)), the largest absolute differences obtain for the parameter values $(g_2, b_2) = (0.75, 0.75)$ on which all posteriors perform best.

Bayes

H^*	Bayes
(0.25, 0.25)	0.97421
(0.25, 0.75)	0.97718
(0.50, 0.50)	1.39253
(0.75, 0.25)	0.99526
(0.75, 0.75)	0.76698

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.97386	0.97395	0.96496
(0.25, 0.75)	0.97692	0.97700	0.96971
(0.50, 0.50)	1.39260	1.39258	1.39611
(0.75, 0.25)	0.99547	0.99542	1.00417
(0.75, 0.75)	0.73453	0.76543	0.67806

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.97386	0.97395	0.96407
(0.25, 0.75)	0.97692	0.97700	0.96888
(0.50, 0.50)	1.39260	1.39258	1.39679
(0.75, 0.25)	0.99547	0.99542	1.00572
(0.75, 0.75)	0.73453	0.76542	0.67690

Our results provide further support for previous works by Glass and co-workers highlighting the performance of PCM ("PCM performs much better than ML and LR in almost all cases and better than OCM for larger sample sizes" [7, p. 191] and "PCM seems preferable to ML as a measure for abductive inference" [8, p. 315]) and the superiority of best explanation posteriors over the Bayesian posterior ("explanationist updating out-performs Bayesian updating on all desirable counts: it takes one, in general, faster to the truth, it minimizes, in general, one's Brier scores, and it leads, in general, to a lower average Brier score" [30, p. 566]), see also [31]. We are thus led to believe that intertwining Bayesian and IBE can be technically fruitful and provide posterior probabilities with desirable characteristics. There seems to be ample scope for further studies of intertwining IBE and Bayesian ideas.

While PCM posteriors outperformed the other posteriors most of the time in our simulations, we also found scenarios in which the PCM posterior performed worst (Table 2). It thus not only the case that the apparent front-runner does not always come first, it sometimes even comes last.

The approach of mixing the best of both worlds by forming Bayes + IBE Posteriors did not have the desired result of a performance increase. Instead, these mixed posteriors closely followed their respective IBE posteriors. We believe that this is because in our simulation environment the absolute values of the explanations, the $\mathcal{E}$, have greater absolute values than the Bayesian part (Equation (2)), since the weighting involves the prior probability of H, which is $1/289$. Table 3 shows that the posteriors defined by Equation (2) and Equation (3) are different.

In every fixed case, all posteriors assign the same hypothesis $H \in \mathbb{H}$ the greatest probability (Tables 6 and 7). In a sense they hence all perform equally well as an inference to the best explanation procedure.

5. Conclusions

Previous works assessed versions of IBE to determine posterior probabilities using synthetic data and found evidence for the superiority of IBE approaches over the Bayesian approach.

We found in our simulations that the best-performing posterior in any given scenario is typically performing worst in other scenarios. This raises the possibility that a general method for obtaining posterior probabilities that is guaranteed to deliver close to optimal performances might not exist. This would be especially detrimental for applications in which there is normally no way to *assess a priori* the performance of posterior probabilities. Post hoc assessments of posteriors may provide some guidance for future applications, but then hindsight is 20/20 and of little use prior to making important decisions.

Table 3
Measure 2: Setting 1: For $(g_2, b_2) = (0.25, 0.25), (0.25, 0.75), (0.5, 0.5), (0.75, 0.75)$ the posterior P_{PCM} outperforms all other posteriors (in terms of minimising d_2). For the other parameter pairs $(g_2, b_2) = (0.75, 0.25)$ the Bayesian posterior, P_{ML} and P_{MPE} and their combinations outperform all other posteriors.
All posteriors but the PCM measures performed worst on the parameters $(g_2, b_2) = (0.5, 0.5)$ and best on the parameters $(g_2, b_2) = (0.75, 0.75)$. The PCM measures performed worst on $(g_2, b_2) = (0.75, 0.25)$ and best on $(g_2, b_2) = (0.75, 0.75)$. While there is in general not much difference in performance between posteriors (for fixed (g_2, b_2)), the largest absolute differences obtain for the parameter values $(g_2, b_2) = (0.25, 0.25)$ on which all posteriors perform worst.

Bayes

H^*	Bayes
(0.25, 0.25)	0.11971
(0.25, 0.75)	0.07728
(0.50, 0.50)	0.05067
(0.75, 0.25)	0.07241
(0.75, 0.75)	0.04521

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.11873	0.11899	0.08899
(0.25, 0.75)	0.07678	0.07694	0.06106
(0.50, 0.50)	0.05029	0.05040	0.03987
(0.75, 0.25)	0.07298	0.07283	0.09453
(0.75, 0.75)	0.04180	0.04503	0.03492

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.11873	0.11898	0.08529
(0.25, 0.75)	0.07678	0.07693	0.05902
(0.50, 0.50)	0.05029	0.05040	0.03895
(0.75, 0.25)	0.07298	0.07284	0.09798
(0.75, 0.75)	0.04180	0.04503	0.03479

From a foundational point of view, the trend that the Bayesian approach, long and widely held to be a gold standard of rationality, is outperformed continues. It is not particularly surprising that there are some instances in which other posteriors outperform the Bayesian posterior, since gerrymandering a posterior that is (very close to) the actual state of the world is always possible and such posteriors tend to outperform all other (including the Bayesian) posteriors. However, the approach that has been shown to outperform Bayesian posterior is based on axiomatic foundations, and hence it does not seem reasonable to explain the performance increase simply as "adhocness".

It is, of course, possible to trace the performance increase to different axiomatic foundations and to assess, on whatever grounds, that the axiomatic foundations for PCM are better than those for the Bayesian approach. We are sceptical that such assessments can shed much light on the observed performance increase. We are instead wondering whether there are, in our view, more illuminating reasons according to which IBE posteriors outperform the Bayesian approach in a (growing) number of simulation environments.

While we have made some progress with evaluating different versions of IBE in this paper, we surely did not (aim to) close the debate. Since i) the quest for determining the best measure of explanation (for a given purposes) on a priori considerations continues [76,77] and further promising measures may be found, ii) we only tested on one (set of) ABMs and it is not necessarily true that comparable results will obtain for other ABMs and iii) none of the evaluations carried out so far have used real-world data, however, arguably, the real challenge is the real world and not a controlled synthetic environment, see [78–80] for comparisons of various forms of Bayesian posteriors (such as imprecise and objective posteriors).

In future work, we plan to assess different versions of IBE with respect to each other and a Bayesian approach on real-world data. Bayesian methods have been indeed suggested and applied for empirical validations and estimations of ABM, e.g., [81–83]. Other possible avenues of future research are i) ABMs with large parameter spaces, which may require techniques to reduce the dimensionality, such as principal component analysis [84,85], ii) the speed with which different updating methods converge to the truth, if they do converge, iii) the search for reasons for the (relative) performances of different updating procedures and iv) the use of background knowledge resulting in non-uniform priors. A non-uniform prior can also shed light on the maximum likelihood posteriors which are equal to the Bayesian posterior in our simulations due to the uniform prior.

Table 4
Measure 1: Setting 2 For $(g_2, b_2) = (0.25, 0.25), (0.25, 0.75), (0.75, 0.25), (0.75, 0.75)$ the posterior P_{PCM} outperforms all other posteriors (in terms of minimising d_1). For the other parameter pair $(g_2, b_2) = (0.5, 0.5)$ the Bayes + P_{PCM} posterior outperforms all other posteriors. This is one of the few instances in which a Bayes + IBE Posterior outperforms both the Bayesian and the IBE posterior (on the level of precision in this table) – other such instances are $(g_2, b_2) = (0.25, 0.25)$ and the ML and MPE posteriors. Since no posteriors performs worse than the Bayesian posterior there is no blue-coloured number.
All posteriors performed worst on the parameters $(g_2, b_2) = (0.5, 0.5)$ and best on the parameters $(g_2, b_2) = (0.75, 0.75)$. While there is in general not much difference in performance between posteriors (for fixed (g_2, b_2)), the largest absolute differences obtain for the parameter values $(g_2, b_2) = (0.75, 0.75)$ on which all posteriors perform best.

Bayes

H^*	Bayes
(0.25, 0.25)	0.98879
(0.25, 0.75)	0.94324
(0.50, 0.50)	1.38050
(0.75, 0.25)	0.93604
(0.75, 0.75)	0.70110

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.98877	0.98877	0.98795
(0.25, 0.75)	0.94232	0.94245	0.93442
(0.50, 0.50)	1.38028	1.38030	1.37963
(0.75, 0.25)	0.93475	0.93498	0.92564
(0.75, 0.75)	0.65103	0.69028	0.60633

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.98877	0.98877	0.98785
(0.25, 0.75)	0.94232	0.94243	0.93412
(0.50, 0.50)	1.38028	1.38030	1.37973
(0.75, 0.25)	0.93475	0.93496	0.92541
(0.75, 0.75)	0.65103	0.69023	0.60544

CRediT authorship contribution statement

Francesco De Pretis: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Aldo Glielmo:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jürgen Landes:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

Francesco De Pretis acknowledges support from the University of Modena and Reggio Emilia, Italy. Jürgen Landes gratefully acknowledges NextGenerationEU funding for the project "Practical Reasoning for Human-Centred Artificial Intelligence" as well as funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 528031869.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Juergen Landes reports financial support was provided by German Research Foundation project number 528031869. Juergen Landes reports financial support was provided by NextGenerationEU via the project Practical Reasoning for Human-Centred Artificial Intelligence.

Acknowledgements

Many thanks to Lorenzo Casini (University of Bologna, Italy), Sébastien Destercke (CNRS, Compiègne, France), Marco Favorito (Banca d'Italia, Italy), Luca Fierro (International Institute for Applied Systems Analysis, Austria) and Borut Trpin (LMU Munich,

Table 5
Measure 2: Setting 2 For $(g_2, b_2) = (0.25, 0.25), (0.25, 0.75), (0.75, 0.25), (0.75, 0.75)$ the posterior P_{PCM} outperforms all other posteriors (in terms of minimising d_2). For the other parameter pair $(g_2, b_2) = (0.5, 0.5)$ the LR posteriors outperform all other posteriors.
All posteriors performed worst on the parameters $(g_2, b_2) = (0.5, 0.5)$ and best on the parameters $(g_2, b_2) = (0.75, 0.75)$. While there is in general not much difference in performance between posteriors (for fixed (g_2, b_2)), the largest absolute differences obtain for the parameter values $(g_2, b_2) = (0.75, 0.75)$ on which all posteriors perform worst.

Bayes

H^*	Bayes
(0.25, 0.25)	0.22311
(0.25, 0.75)	0.17717
(0.50, 0.50)	0.27547
(0.75, 0.25)	0.19708
(0.75, 0.75)	0.14380

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.22226	0.22238	0.18720
(0.25, 0.75)	0.17136	0.17215	0.11534
(0.50, 0.50)	0.27529	0.27531	0.29453
(0.75, 0.25)	0.19313	0.19385	0.16403
(0.75, 0.75)	0.13401	0.14167	0.12580

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	0.22226	0.22228	0.18131
(0.25, 0.75)	0.17136	0.17201	0.11315
(0.50, 0.50)	0.27529	0.27531	0.29668
(0.75, 0.25)	0.19313	0.19378	0.16326
(0.75, 0.75)	0.13401	0.14166	0.12564

Table 6
Setting 1 Hypotheses with largest posterior probabilities, in terms of (g_2, b_2), and their respective probability.

Bayes

H^*	Bayes
(0.25, 0.25)	(0.10, 0.25), 0.04901
(0.25, 0.75)	(0.20, 0.80), 0.03906
(0.50, 0.50)	(0.20, 0.70), 0.03498
(0.75, 0.25)	(0.10, 0.45), 0.03751
(0.75, 0.75)	(0.75, 0.75), 0.32948

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	(0.10, 0.25), 0.05021	(0.10, 0.25), 0.04989	(0.10, 0.25), 0.08889
(0.25, 0.75)	(0.20, 0.80), 0.03960	(0.20, 0.80), 0.03943	(0.20, 0.80), 0.05701
(0.50, 0.50)	(0.20, 0.70), 0.03542	(0.20, 0.70), 0.03529	(0.20, 0.70), 0.05103
(0.75, 0.25)	(0.10, 0.45), 0.03807	(0.10, 0.45), 0.03793	(0.10, 0.45), 0.05795
(0.75, 0.75)	(0.75, 0.75), 0.36487	(0.75, 0.75), 0.33125	(0.75, 0.75), 0.43760

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	(0.10, 0.25), 0.05021	(0.10, 0.25), 0.04990	(0.10, 0.25), 0.09431
(0.25, 0.75)	(0.20, 0.80), 0.03960	(0.20, 0.80), 0.03944	(0.20, 0.80), 0.05943
(0.50, 0.50)	(0.20, 0.70), 0.03542	(0.20, 0.70), 0.03529	(0.20, 0.70), 0.05346
(0.75, 0.25)	(0.10, 0.45), 0.03807	(0.10, 0.45), 0.03793	(0.10, 0.45), 0.06101
(0.75, 0.75)	(0.75, 0.75), 0.36487	(0.75, 0.75), 0.33126	(0.75, 0.75), 0.43911

Table 7
Setting 2 Hypotheses with largest posterior probabilities and their respective probability.

Bayes

H^*	Bayes
(0.25, 0.25)	(0.25, 0.20), 0.02232
(0.25, 0.75)	(0.35, 0.65), 0.09581
(0.50, 0.50)	(0.20, 0.75), 0.09398
(0.75, 0.25)	(0.65, 0.30), 0.11812
(0.75, 0.75)	(0.75, 0.75), 0.41796

Bayes + IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	(0.25, 0.20), 0.02242	(0.25, 0.20), 0.02240	(0.25, 0.20), 0.02697
(0.25, 0.75)	(0.35, 0.65), 0.09767	(0.35, 0.65), 0.09741	(0.35, 0.65), 0.11748
(0.50, 0.50)	(0.20, 0.75), 0.09673	(0.20, 0.75), 0.09634	(0.20, 0.75), 0.13049
(0.75, 0.25)	(0.65, 0.30), 0.12309	(0.65, 0.30), 0.12216	(0.65, 0.30), 0.17245
(0.75, 0.75)	(0.75, 0.75), 0.47083	(0.75, 0.75), 0.42974	(0.75, 0.75), 0.53227

IBE Posteriors

H^*	LR	OCM	PCM
(0.25, 0.25)	(0.25, 0.20), 0.02242	(0.25, 0.20), 0.02242	(0.25, 0.20), 0.02787
(0.25, 0.75)	(0.35, 0.65), 0.09767	(0.35, 0.65), 0.09746	(0.35, 0.65), 0.11845
(0.50, 0.50)	(0.20, 0.75), 0.09673	(0.20, 0.75), 0.09641	(0.20, 0.75), 0.13238
(0.75, 0.25)	(0.65, 0.30), 0.12310	(0.65, 0.30), 0.12224	(0.65, 0.30), 0.17480
(0.75, 0.75)	(0.75, 0.75), 0.47083	(0.75, 0.75), 0.42980	(0.75, 0.75), 0.53348

Germany) for helpful discussions.
The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the official policy or position of Banca d'Italia.

Appendix A

Table A.1
Step by step pseudo code detailing the implementation of our simulations.

Step	Task	Description
1	Initialise Model Parameters	- Define number of strategies $S = 4$. - Initialise parameters g_s and b_s for all strategies s. - Set global parameters R, β, σ, using pre-defined values: $(R, \beta, \sigma) = (1.01, 10, 0.04)$.
2	Setting Simulation Environment	- Parameters: $g_2, b_2 \in [0.1, 0.9]$ generate set of hypothesis, $\mathbb{H}$. - Initialise parameter values from Table 1 for Setting 1 and Setting 2. - Do all of the following for Setting 1 and Setting 2.
3	Compute Time Series	- For every hypothesis $H \in \mathbb{H}$, generate 10,000 time series, 200 time steps each.
4	Compute First and Second Moments	- For every time series, compute the average of asset prices (first moment) over 200 time steps. - For every time series, compute the variance (second moment) over 200 time steps.
5	Tessellation and Likelihoods	- Cover the space spanned by first and second moments with a hexagonal tessellation. - For every hypothesis H and every tile of the tessellation, calculate the frequency of time series with first and second moments being within this tile.
6	Evidence	- Define five actual hypotheses H^* given by: $(0.25, 0.25), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25)$ and $(0.75, 0.75)$. - Do all of the following for all H^*. - For every H^* compute the tile E^* with greatest likelihood.
7	Posteriors	- For every hypothesis $H \in \mathbb{H}$, compute the Bayesian posterior $P_{\text{Bayes}}(H\|E^*)$. - For every hypothesis $H \in \mathbb{H}$, compute the IBE posterior $P_{\text{IBE}}(H\|E^*)$. - For every hypothesis $H \in \mathbb{H}$, compute the Bayes + IBE posterior $P_{\text{Bayes+IBE}}(H\|E^*)$.
8	Parameter Approximation (Measure 1)	- For all posteriors, calculate the Euclidean distance, d_1, to the true posterior given by $P(H^*) = 1$.
9	Forecast Accuracy (Measure 2)	- For all posteriors, calculate the probability of every tile, $P_{\text{post}}(E)$. - For all posteriors, calculate the Euclidean difference, d_2, between the $P_{\text{post}}(E)$ and the $P(E\|H^*)$.
10	Greatest Probabilities	- For every posterior, compute the tile with the greatest posterior probability and compute this posterior probability.

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijar.2025.109388.

Data availability

All data used in this study is publicly available at the specified websites. Computer code used in this study is either publicly available or attached.

References

[1] P. Lipton, Inference to the Best Explanation, 2nd edition, Routledge, London, 2004.
[2] I. Douven, Optimizing group learning: an evolutionary computing approach, Artif. Intell. 275 (2019) 235–251, https://doi.org/10.1016/j.artint.2019.06.002.
[3] I. Douven, Abduction: theory and evidence, in: L. Magnani (Ed.), Handbook of Abductive Cognition, Springer International Publishing, Cham, 2022, pp. 1–29, Ch. 61.
[4] V.F. Hendricks, J. Faye, Abducting explanation, in: L. Magnani, N.J. Nersessian, P. Thagard (Eds.), Model-Based Reasoning in Scientific Discovery, Springer, Boston, MA, 1999, pp. 271–292.
[5] D.H. Glass, Coherence measures and inference to the best explanation, Synthese 157 (3) (2006) 275–296, https://doi.org/10.1007/s11229-006-9055-7.
[6] D.H. Glass, Inference to the best explanation: does it track truth?, Synthese 185 (3) (2010) 411–427, https://doi.org/10.1007/s11229-010-9829-9.
[7] D.H. Glass, An evaluation of probabilistic approaches to inference to the best explanation, Int. J. Approx. Reason. 103 (2018) 184–194, https://doi.org/10.1016/j.ijar.2018.09.004.
[8] J.-D. Huang, D.H. Glass, M. McCartney, A comparison of explanatory measures in abductive inference, in: M.J. Lesot, S. Vieira, M.Z. Reformat, J.P. Carvalho, A. Wilbik, B. Bouchon-Meunier, R.R. Yager (Eds.), Information Processing and Management of Uncertainty in Knowledge-Based Systems. IPMU 2020, Communications in Computer and Information Science, Springer International Publishing, 2020, pp. 304–317.
[9] B. Trpin, M. Pellert, Inference to the best explanation in uncertain evidential situations, Br. J. Philos. Sci. 70 (4) (2019) 977–1001, https://doi.org/10.1093/bjps/axy027.
[10] V. Crupi, K. Tentori, A second look at the logic of explanatory power (with two novel representation theorems), Philos. Sci. 79 (3) (2012) 365–385, https://doi.org/10.1086/666063.
[11] S. Dragulinescu, Inference to the best explanation as a theory for the quality of mechanistic evidence in medicine, Eur. J. Philos. Sci. 7 (2) (2016) 353–372, https://doi.org/10.1007/s13194-016-0165-x.
[12] F. Dellsén, Abductive Reasoning in Science, Cambridge University Press, Cambridge, 2024.
[13] J.N. Schupbach, On the logical structure of best explanations, Philos. Sci. 90 (5) (2023) 1150–1160, https://doi.org/10.1017/psa.2023.45.
[14] I. Douven, The Art of Abduction, MIT Press, Cambridge, 2022.
[15] I. Douven, Abduction, in: E.N. Zalta (Ed.), Stanford Encyclopedia of Philosophy, Summer 2021 Edition, Metaphysics Research Lab, Stanford University, 2021, https://plato.stanford.edu/archives/sum2021/entries/abduction.
[16] M. Tesic, B. Eva, S. Hartmann, Confirmation by explanation: a Bayesian justification of IBE unpublished manuscript, https://doi.org/10.5282/ubm/epub.41934, 2017.
[17] D. Šešelja, C. Straßer, Abstract argumentation and explanation applied to scientific debates, Synthese 190 (12) (2013) 2195–2217, https://doi.org/10.1007/s11229-011-9964-y.
[18] L. Amgoud, Explaining black-box classifiers: properties and functions, Int. J. Approx. Reason. 155 (2023) 40–65, https://doi.org/10.1016/j.ijar.2023.01.004.
[19] J. Weisberg, Locating IBE in the Bayesian framework, Synthese 167 (1) (2008) 125–143, https://doi.org/10.1007/s11229-008-9305-y.
[20] C.M. Glen, M.L. Kemp, E.O. Voit, Agent-based modeling of morphogenetic systems: advantages and challenges, PLoS Comput. Biol. 15 (3) (2019) e1006577, https://doi.org/10.1371/journal.pcbi.1006577.
[21] E. Hunter, B. Mac Namee, J.D. Kelleher, A taxonomy for agent-based models in human infectious disease epidemiology, J. Artif. Soc. Soc. Simul. 20 (3) (2017), https://doi.org/10.18564/jasss.3414.
[22] R.L. Axtell, J.D. Farmer, Agent-based modeling in economics and finance: past, present, and future, J. Econ. Lit. (2022) 1–101, https://www.aeaweb.org/articles?id=10.1257/jel.20221319.
[23] C.O. Retzlaff, M. Ziefle, A. Calero Valdez, The history of agent-based modeling in the social sciences, in: V.G. Duffy (Ed.), Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Body, Motion and Behavior. HCII 2021, in: Lecture Notes in Computer Science, Springer, Cham, 2021, pp. 304–319.
[24] F. Lamperti, A. Roventini, A. Sani, Agent-based model calibration using machine learning surrogates, J. Econ. Dyn. Control 90 (2018) 366–389, https://doi.org/10.1016/j.jedc.2018.03.011.
[25] D. Platt, A comparison of economic agent-based model calibration methods, J. Econ. Dyn. Control 113 (2020) 103859, https://doi.org/10.1016/j.jedc.2020.103859.
[26] A. Glielmo, M. Favorito, D. Chanda, D. Delli Gatti, Reinforcement learning for combining search methods in the calibration of economic ABMs, in: Proceedings of the Fourth ACM International Conference on AI in Finance, Association for Computing Machinery, New York, 2023, pp. 305–313.
[27] J. Grazzini, M.G. Richiardi, M. Tsionas, Bayesian estimation of agent-based models, J. Econ. Dyn. Control 77 (2017) 26–47, https://doi.org/10.1016/j.jedc.2017.01.014.
[28] J. Dyer, P. Cannon, J.D. Farmer, S.M. Schmon, Black-box Bayesian inference for agent-based models, J. Econ. Dyn. Control 161 (2024) 104827, https://doi.org/10.1016/j.jedc.2024.104827.
[29] I. Douven, Inference to the best explanation, Dutch books, and inaccuracy minimisation, Philos. Q. 63 (252) (2013) 428–444, https://doi.org/10.1111/1467-9213.12032.
[30] I. Douven, S. Wenmackers, Inference to the best explanation versus Bayes's rule in a social setting, Br. J. Philos. Sci. 68 (2) (2017) 535–570, https://doi.org/10.1093/bjps/axv025.
[31] I. Douven, The ecological rationality of explanatory reasoning, Stud. Hist. Phil. Sci. A 79 (2020) 1–14, https://doi.org/10.1016/j.shpsa.2019.06.004.
[32] A. Reutlinger, Explanation beyond causation? New directions in the philosophy of scientific explanation, Philos. Compass 12 (2) (2017) e12395, https://doi.org/10.1111/phc3.12395.
[33] A. Brenner, A.-S. Maurin, A. Skiles, R. Stenwall, N. Thompson, Metaphysical explanation, in: E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Winter 2021 Edition, Metaphysics Research Lab, Stanford University, 2021, https://plato.stanford.edu/archives/win2021/entries/metaphysical-explanation.
[34] P. Mancosu, F. Poggiolesi, C. Pincock, Mathematical explanation, in: E.N. Zalta, U. Nodelman (Eds.), The Stanford Encyclopedia of Philosophy, Fall, 2023, 2023 Edition, Metaphysics Research Lab, Stanford University, 2023, https://plato.stanford.edu/archives/win2021/entries/metaphysical-explanation.

[35] J. Woodward, L. Ross, Scientific explanation, in: E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Summer 2021 Edition, Metaphysics Research Lab, Stanford University, 2021, https://plato.stanford.edu/archives/sum2021/entries/scientific-explanation.
[36] J. Borrego-Díaz, J. Galán-Páez, Explainable artificial intelligence in data science, Minds Mach. 32 (2022) 485–531, https://doi.org/10.1007/s11023-022-09603-z.
[37] D.S. Watson, L. Floridi, The explanation game: a formal framework for interpretable machine learning, Synthese 198 (10) (2021) 9211–9242, https://doi.org/10.1007/s11229-020-02629-9.
[38] G. Ras, N. Xie, M.V. Gerven, D. Doran, Explainable deep learning: a field guide for the uninitiated, J. Artif. Intell. Res. 73 (2022) 329–397, https://doi.org/10.1613/jair.1.13200.
[39] B. Trpin, Affirming the explanandum, analysis, Ahead of print, https://doi.org/10.1093/analys/anae003, 2025.
[40] H. Hosni, J. Landes, Logical perspectives on foundations of probability, Open Math. 21 (1) (2023) 20220598, https://doi.org/10.1515/math-2022-0598.
[41] R.T. Stewart, M. Nielsen, Another approach to consensus and maximally informed opinions with increasing evidence, Philos. Sci. 86 (2) (2019) 236–254, https://doi.org/10.1086/701954.
[42] M.J. Schervish, A general method for comparing probability assessors, Ann. Stat. 17 (4) (1989) 1856–1879, https://doi.org/10.1214/aos/1176347398.
[43] L. Tesfatsion, Agent-based computational economics: modeling economies as complex adaptive systems, Inf. Sci. 149 (4) (2003) 262–268, https://doi.org/10.1016/S0020-0255(02)00280-3.
[44] T. Lux, M. Marchesi, Scaling and criticality in a stochastic multi-agent model of a financial market, Nature 397 (6719) (1999) 498–500, https://doi.org/10.1038/17290.
[45] W.B. Arthur, The Economy as an Evolving Complex System II, CRC Press, 2018.
[46] G. Dosi, M. Napoletano, A. Roventini, J.E. Stiglitz, T. Treibich, Rational heuristics? Expectations and behaviors in evolving economies with heterogeneous interacting agents, Econ. Inq. 58 (3) (2020) 1487–1516, https://doi.org/10.1111/ecin.12897.
[47] M. Radzvilas, F. De Pretis, W. Peden, D. Tortoli, B. Osimani, Incentives for research effort: an evolutionary model of publication markets with double-blind and open review, Comput. Econ. 61 (4) (2023) 1433–1476, https://doi.org/10.1007/s10614-022-10250-w.
[48] J.D. Farmer, D. Foley, The economy needs agent-based modelling, Nature 460 (7256) (2009) 685–686, https://doi.org/10.1038/460685a.
[49] L. Hamill, N. Gilbert, Agent-Based Modelling in Economics, Wiley, 2015.
[50] M. Napoletano, E. Guerci, N. Hanaki, Recent advances in financial networks and agent-based model validation, J. Econ. Interact. Coord. 13 (1) (2018) 1–7, https://doi.org/10.1007/s11403-018-0221-z.
[51] J.-P. Bouchaud, Agent-based models for market impact and volatility, in: C. Hommes, B. LeBaron (Eds.), Handbook of Computational Economics, vol. 4, Elsevier, 2018, pp. 393–436, Ch. 7.
[52] W.A. Brock, C.H. Hommes, Heterogeneous beliefs and routes to chaos in a simple asset pricing model, J. Econ. Dyn. Control 22 (8–9) (1998) 1235–1274, https://doi.org/10.1016/s0165-1889(98)00011-6.
[53] R. Jamali, W. Vermeiren, S. Lazarova-Molnar, Data-driven agent-based modeling: experimenting with the Schelling's model, Proc. Comput. Sci. 238 (2024) 298–305, https://doi.org/10.1016/j.procs.2024.06.028.
[54] J.N. Schupbach, D.H. Glass, Hypothesis competition beyond mutual exclusivity, Philos. Sci. 84 (5) (2017) 810–824, https://doi.org/10.1086/693928.
[55] D.H. Glass, Competing hypotheses and abductive inference, Ann. Math. Artif. Intell. 89 (1–2) (2021) 161–178, https://doi.org/10.1007/s10472-019-09630-0.
[56] J.N. Schupbach, D.H. Glass, Conjunctive Explanations: The Nature, Epistemology, and Psychology of Explanatory Multiplicity, Routledge, New York, 2023.
[57] D.H. Glass, J.N. Schupbach, Conjunctive explanations: when are two explanations better than one?, Synthese 204 (2) (2024) 49, https://doi.org/10.1007/s11229-024-04683-z.
[58] P. Ylikoski, J. Kuorikoski, Dissecting explanatory power, Philos. Stud. 148 (2) (2010) 201–219, https://doi.org/10.1007/s11098-008-9324-z.
[59] V. Crupi, K. Tentori, State of the field: measuring information and confirmation, Stud. Hist. Phil. Sci. A 47 (2014) 81–90, https://doi.org/10.1016/j.shpsa.2014.05.002.
[60] J.N. Schupbach, J. Sprenger, The logic of explanatory power, Philos. Sci. 78 (1) (2011) 105–127, https://doi.org/10.1086/658111.
[61] I.J. Good, Weight of evidence, corroboration, explanatory power, information and the utility of experiments, J. R. Stat. Soc., Ser. B, Stat. Methodol. 22 (2) (1960) 319–331, https://doi.org/10.1111/j.2517-6161.1960.tb00378.x.
[62] T. McGrew, Confirmation, heuristics, and explanatory reasoning, Br. J. Philos. Sci. 54 (4) (2003) 553–567, https://doi.org/10.1093/bjps/54.4.553.
[63] D.H. Glass Coherence Explanation, Hypothesis Selection, Br. J. Philos. Sci. 72 (1) (2021) 1–26, https://doi.org/10.1093/bjps/axy063.
[64] C. Glymour, Probability and the explanatory virtues, Br. J. Philos. Sci. 66 (3) (2015) 591–604, https://doi.org/10.1093/bjps/axt051.
[65] J. Landes, Probabilism, entropies and strictly proper scoring rules, Int. J. Approx. Reason. 63 (2015) 1–21, https://doi.org/10.1016/j.ijar.2015.05.007.
[66] J. Williamson, in: Defence of Objective Bayesianism, Oxford University Press, Oxford, 2010.
[67] C. Hommes, Heterogeneous agent models: two simple examples, in: M. Lines (Ed.), Nonlinear Dynamical Systems in Economics, Springer, 2005, pp. 131–164, Ch. 5.
[68] C.H. Hommes, Modeling the stylized facts in finance through simple nonlinear adaptive systems, Proc. Natl. Acad. Sci. 99 (suppl_3) (2002) 7221–7228, https://doi.org/10.1073/pnas.082080399.
[69] M. Benedetti, G. Catapano, F.D. Sclavis, M. Favorito, A. Glielmo, D. Magnanimi, A. Muci, Black-it: a ready-to-use and easy-to-extend calibration kit for agent-based models, J. Open Sour. Softw. 7 (79) (2022) 4622, https://doi.org/10.21105/joss.04622.
[70] G.W. Brier, Verification of forecasts expressed in terms of probability, Mon. Weather Rev. 78 (1) (1950) 1–3, https://doi.org/10.1175/1520-0493(1950)078%3C0001:VOFEIT%3E2.0.CO;2.
[71] J.-C. Billaut, D. Bouyssou, P. Vincke, Should you believe in the Shanghai ranking? An MCDM view, Scientometrics 84 (2010) 237–263, https://doi.org/10.1007/s11192-009-0115-x.
[72] J.E. Bickel, Some comparisons among quadratic, spherical, and logarithmic scoring rules, Decis. Anal. 4 (2) (2007) 49–65, https://doi.org/10.1287/deca.1070.0089.
[73] D.V. Lindley, Scoring rules and the inevitability of probability, Int. Stat. Rev. / Rev. Int. Stat. 50 (1) (1982) 1–11, https://doi.org/10.2307/1402448.
[74] J. Predd, R. Seiringer, E. Lieb, D. Osherson, H. Poor, S. Kulkarni, Probabilistic coherence and proper scoring rules, IEEE Trans. Inf. Theory 55 (10) (2009) 4786–4792, https://doi.org/10.1109/TIT.2009.2027573.
[75] L.J. Savage, Elicitation of personal probabilities and expectations, J. Am. Stat. Assoc. 66 (336) (1971) 783–801, https://doi.org/10.1080/01621459.1971.10482346.
[76] D.H. Glass, How good is an explanation?, Synthese 201 (2) (2023) 53, https://doi.org/10.1007/s11229-022-04025-x.
[77] S. Hartmann, B. Trpin, Conjunctive explanations: a coherentist appraisal, in: J.N. Schupbach, D.H. Glass (Eds.), Conjunctive Explanations: The Nature, Epistemology, and Psychology of Explanatory Multiplicity, Routledge, New York, 2023, pp. 111–142.
[78] M. Radzvilas, W. Peden, F. De Pretis, Making decisions with evidential probability and objective bayesian calibration inductive logics, Int. J. Approx. Reason. 162 (2023) 109030, https://doi.org/10.1016/j.ijar.2023.109030.
[79] M. Radzvilas, W. Peden, F. De Pretis, The ambiguity dilemma for imprecise bayesians, Ahead of print, Br. J. Philos. Sci. (2024), https://doi.org/10.1086/729618.
[80] M. Radzvilas, W. Peden, D. Tortoli, F. De Pretis, A comparison of imprecise bayesianism and Dempster-Shafer theory for automated decisions under ambiguity, Ahead of print, J. Log. Comput. (2024), https://doi.org/10.1093/logcom/exae069.
[81] G. Fagiolo, M. Guerini, F. Lamperti, A. Moneta, A. Roventini, Validation of agent-based models in economics and finance, in: C. Beisbart, N.J. Saam (Eds.), Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives, Springer International Publishing, 2019, pp. 763–787.

[82] J. Grazzini, M. Richiardi, Estimation of ergodic agent-based models by simulated minimum distance, J. Econ. Dyn. Control 51 (2015) 148–165, https://doi.org/10.1016/j.jedc.2014.10.006.

[83] K.B. Korb, N. Geard, A. Dorin, A Bayesian approach to the validation of agent-based models, in: A. Tolk (Ed.), Ontology, Epistemology, and Teleology for Modeling and Simulation: Philosophical Foundations for Intelligent M&S Applications, Springer, Berlin, 2013, pp. 255–269.

[84] P. Hünermund, B. Louw, I. Caspi, Double machine learning and automated confounder selection: a cautionary tale, J. Causal Inference 11 (1) (2023), https://doi.org/10.1515/jci-2022-0078.

[85] F. Zayas-Gato, Á. Michelena, H. Quintián, E. Jove, J.-L. Casteleiro-Roca, P. Leitão, J.L. Calvo-Rolle, A novel method for anomaly detection using beta Hebbian learning and principal component analysis, Log. J. IGPL 31 (2) (2023) 390–399, https://doi.org/10.1093/jigpal/jzac026.