Eur. Phys. J. Special Topics **229**, 1645–1659 (2020)

https://doi.org/10.1140/epjst/e2020-900262-y

THE EUROPEAN PHYSICAL JOURNAL
SPECIAL TOPICS

Regular Article

What is an explanation? Statistical physics and economics

Sanjay G. Reddy[a]

Department of Economics, The New School for Social Research, New York, NY, USA

Received 3 November 2019 / Received in final form 29 February 2020
Published online 7 July 2020

Abstract. There is a long history of statistical physics being looked to as a source of inspiration for the social sciences. There has been growing interest in the field. However, the bridge between the two has been imperfect. One way to understand why is that there are differences between their objects of explanation. While approaches derived from statistical physics may offer a useful tool in specific settings, it is vital to recognize when the insights that derive from them may be limited as a result of differences between the social and the physical world. We argue that methods from statistical physics can help in providing explanations of economic outcomes only if they include adequate attention to processes and moreover characterize these in an appropriate way, recognizing their specifically social character.

1 What is an explanation?

Can statistical physics approaches help economics? This question might be answered according to different criteria, depending on what one sees as the goals of the discipline. These goals include description, prediction and explanation (see [47–49]).

In this paper, we address the extent to which statistical physics approaches in economics are likely to help to *explain* economic processes and outcomes. An explanation requires relating statements that are the objects of explanation – what is to be explained (or the *explanandum*) – to other statements which are the explanations – that which does the explaining (or the *explanans*). The relations that play this role can be empirical or logical in nature, or both, but must help to answer "why" and not merely to describe "what" [26]. For example, an explanation might involve the idea that an empirical process has caused a set of circumstances, defining a state A, to give rise, over time, to another state B, or it might involve the demonstration that a set of conditions C logically implies another set of conditions D. In either case, explanation as we use the term here, involves an effort to understand why something is the case by relating it to other things. Explanation is therefore treated here as a distinct concept from both description and prediction, which respectively centre on characterization of the present or the past, or anticipation of the future, but which need not require answering why, for instance by relating present to past[1]. Since

[a] e-mail: reddysanjayg@gmail.com

[1] Description can include the description of processes and causes, and not merely of states. While it can encompass the relation of states to one another, it need not. In order to avoid confusion, we use the term narrowly to focus on description of states.

explanation is concerned with *why* things happen, it also often involves a description of the mechanisms involved in change, or of *how* things happen. This is a description of processes and not merely of the likelihood of a change from one state to another state (as for instance in a Markov chain). Explanation is important for science because it is required if we are to make sense of why things are the way they are. It may also aid us in employing knowledge fruitfully so as to intervene in in the world effectively (since for that purpose, understanding the pathways through which causes have effects may often be as important as knowing what those effects are)[2]. We take explanation to be a central goal for science in general, including economics and the social sciences[3].

The concept of natural selection provides an apposite example. It provides a method of explanation insofar as it helps us to understand the process by which specific organisms came into existence. It also influences description (by shaping our understanding of the relations between species and therefore the system of classification applied to them). However, while it may have strong predictive power in specific environments (e.g. when applied to the evolution of colonies of bacteria) it is of limited value in predicting the specific outcomes that will arise in many others [24]. Similarly, understanding the origins of the French revolution may or may not provide a guide as to whether other revolutions will occur. A good explanation may not have corresponding predictive power. Obversely, a framework that has predictive power may not provide very much by way of an explanation. For example, an agnostic machine learning algorithm might succeed in predicting a particular phenomenon reliably, without necessarily providing an externally transmissible understanding as to the *reasons* why it arises[4]. It is questionable whether such an approach, however successful, can be thought of as having provided an explanation and, as such, contributed to understanding. While both prediction and explanation may be desirable, they are not always jointly attained, as the example of natural selection illustrates. A scientific theory ought to possess some power of explanation, whether or not it possesses power of prediction. Do statistical physics approaches in economics do so? We shall argue that they possess limited explanatory capability, for the social sciences generally and for economics specifically, at least in their current form.

2 Typology of methods

In what ways do statistical physics techniques in economics relate to methods already in use in the field?

One way to approach this question is to begin with a typology of methods in statistical physics and their analogues in economics. We follow Reif [42] to generate such a typology. Reif provides a four-fold classification of approaches in the field.

[2] See e.g. Deaton and Cartwright [6]. Moving from "in-sample" to "out-of-sample" prediction can require a knowledge of causal processes, for instance, in order to make suitable extrapolations (consider the case of the movement of a rocket in the presence of various bodies generating a gravitational field in an unfamiliar environment – what those bodies are and where they are will permit us to work out the trajectory only if we know the relevant laws).

[3] There are of course debates as to how to understand *exactly* what explanation is in the context of science [61] but these are not directly relevant to our current purpose.

[4] There is some discussion presently on the question of whether, to what extent and under what conditions machine learning methods can generate interpretable results. See e.g. Murdoch et al. [38].

The first set of approaches (or "classical thermodynamics") seeks to identify the "relationships existing between the macroscopic parameters" of "a system in equilibrium". Economics contains a corresponding tradition of studying the relation between macroeconomic aggregates (e.g. much of historical Keynesian as well as contemporary "Post-Keynesian" economic research has been concerned with these relations, both logical and empirical).

The second set of approaches (or "statistical mechanics") seeks to identify statements about "a system in equilibrium" based "on the microscopic properties of the particles in the system and on the laws of mechanics governing their behavior". This is a terrain to which many recent approaches to economic analysis inspired by statistical physics, in particular the "econophysics" literature (see e.g. Dragulescu and Yakovenko [11], the articles surveyed in Gallegati [20], or Shaikh [50–53]) have belonged. Methods falling under this heading focus on identifying systemic economic properties or regularities (e.g. concerning distributions of income or wealth, by analogy with distributions of physical properties such as energy) that are deemed likely to arise as a consequence of the application of statistical laws to the behaviour of individual agents (by analogy with physical particles). Agents are generally not assumed to maximize, since their statistical propensities suffice to drive the system.[5]. The concept of equilibrium that is employed refers to the stability over time of the relevant systemic features rather than to the idea, as in conventional "neoclassical" economics that agents undertake perfectly maximizing behaviors (consumer or producer "equilibrium").

The third set of approaches identified by Reif relates to systems not in equilibrium, for which, we are told, it is possible to make only "very general statements" based on the study of irreversible processes (e.g. "irreversible thermodynamics"). This is the domain for which analogies with economics are perhaps most elusive. The economic "equivalent" might be taken to involve systems which do not possess invariant systemic properties because of the centrality to them of "disequilibrium" processes. There is a literature which refers to such ideas (e.g. within the Keynesian and post-Keynesian traditions as well as in Marxian economics and in ecological economics) and which gives them considerable and even central importance, but the disequilibrium vision has been paid much less attention in the economics discipline than has the equilibrium one. In economics as in statistical physics, the formal modelling of disequilibrium processes is thought to be challenging, and no commonly accepted model exists, although it has been argued to be central to understanding economic phenomena such as unemployment (see e.g. [34]).

The fourth set of approaches identified by Reif relates to the "study in detail" of the "interactions of all the particles in a system" and the calculation of "parameters of macroscopic significance". While such a method is in principle applicable to systems not in equilibrium ("kinetic theory"), it is difficult to implement as it may need a detailed understanding of dynamics and interactions at the level of particles as well as enormous information and computing power. The economic analogy to such an approach is that of an economic model grounded in "micro-foundations" but not necessarily assuming equilibrium. Although the micro-foundations that are most frequently assumed in economics involve "maximizing" agents (firms or consumers) in market or strategic equilibrium, behaviours of any kind can in principle be considered (see e.g. [2,54]) and equilibrium is not always assumed (for instance, in agent-based models, which study the evolution of a system over time, without convergence to an equilibrium being a necessary property). In practice, mainstream economic models built on "micro-foundations", which have been greatly influential in recent

[5] There are exceptions, such as Venkatasubramanian [58], which applies a maximizing model.

decades, have involved extreme simplifications in order to generate tractability, such as the assumption that the agents are all of one type (a "representative agent") and therefore behave identically in relevant respects. They moreover assume equilibrium. In contrast, heterogeneous agent models permitting the study of more complex interactions have been relatively little and only recently explored, in part because of the challenges of tractability and interpretability that they encounter. It was not clear to us whether there are any attempts to apply statistical physics to economics that fall within the category of studying in detail the interactions of heterogeneous agents.

As noted, the statistical physics approaches applied to economics in recent years have largely fallen under the second heading. Another category, inspired by Jaynes [29] and earlier writings) and represented by the recent work of the "New School" group discussed further below, seems difficult to classify in terms of the "traditional" typology provided by Reif[6]. It depends on the idea of entropy maximization as an epistemic postulate, and may or may not be mapped onto a portrait of "microscopic properties of the particles in the system". Indeed, one strand in the literature has emphasized the ability of an approach based on entropy maximization to support a holistic description of systemic dynamics that is not reducible to an account of actions of individuals (see dos Santos in [9])[7].

3 Two strands in the contemporary revival of statistical mechanics in economics

Although there is an earlier history of interest in the potential application of statistical mechanics approaches in economics (see e.g. Samuelson [45], Georgescu-Roegen [21], and for a more general description of the relevant history [40,44]) there has been a revival of such approaches in recent years. The contemporary revival has consisted of two main strands so far, the first focusing on the relation between statistical processes and predicted outcomes, and the second focusing on the relation between epistemic postulates and predicted outcomes. We refer to the first as the processual view and the second as the epistemic view.

The first contribution involves the idea of equilibrium as stability of a statistical distribution. The recent theoretical contributions have shown that such a stable distribution can be predicted to result if specific underlying micro-processes are assumed. Crucially, and in a departure from "standard" microeconomics, optimization by individual agents is not required, and ongoing random transitions between states can suffice to generate and maintain the required stability (see e.g. [11], Shaikh [53]).

The second contribution involves the idea that entropy maximization[8] provides a methodologically privileged approach to predicting what distributions will arise.

[6] For a representative analysis see Semieniuk and Scharfenaker [46]. For programmatic statements see e.g. Foley [15], Foley and Scharfenaker [16] and dos Santos [8,9].

[7] Some of the critics of "methodological individualism" in social science, notably Lukes [33] and Elias [12], have favoured what might be described as a compatibilist perspective that sees a simultaneous role both for social and individual level descriptions.

[8] We recognize that entropy has been thought of in more than one way, and that there may be definite reasons to prefer one concept of entropy over another (see e.g. [27]). We do not, however, assume a resolution to this question as it is unimportant for our purpose, which is to discuss streams of thinking about how concepts from statistical physics can be best applied to the social sciences. For this general purpose, it is the concept of entropy and not the specific conception which is relevant.

In this framework, the specific derivations arrived at depend on the constraints on systemic or individual behavior that are assumed. These constraints correspond to specific assumptions as to what is known, and determine the results of the exercise (see e.g. Foley [15]).

In both cases, there has been a claim of "unreasonable effectiveness" of statistical mechanics approaches, which have been argued to reproduce observed distributions to a seemingly uncanny degree (see e.g. [11], Shaikh [53] and [46]).

The processual view presupposes that an equilibrium distribution is attained as the asymptotic outcome of some process bringing about transitions between attainable states. Irrespective of the precision, or the accuracy, with which specific outcomes are predicted, the process must be specified in order to provide the causal insight that is needed for there to be an explanation (for instance, through an appeal to wage and profit dynamics giving rise to a Fokker–Planck equation[9]).

The epistemic view is in contrast agnostic about the processes involved in giving rise to a predicted outcome, or to put it differently, is not especially concerned with providing an account of such processes. Instead, it focuses on the idea that any given realization of world (or attainable state) among those characterized by a given set of constraints (corresponding to the analyst's state of knowledge) is drawn from the "maximum entropy" distribution that assigns equal probability to each of them (Jaynes [29] provides a clear statement of what we understand by the epistemic view). If the world for which predictions are being generated is treated as having been drawn from this distribution of attainable states, then it may become possible to identify "macroscopic" properties which are in a close range of one another for a very large share of the attainable states. The epistemic view relies on the assumption that we have adequately characterized the constraints, which summarize all relevant information, and which characterize the attainable states. It does not make explicit assumptions about the process that gives rise to one state rather than another, and does not therefore rely on assumptions such as that of ergodicity which characterize the dynamics involved in the evolution of a system, but it arrives at definite conclusions, by assigning equal probability to all states which are deemed consistent with a given state of knowledge (the Principle of Indifference elaborated by Jaynes [29] and drawing on Keynes [31]). The epistemic view treats any particular world – such as our own – as being "drawn" from the maximum entropy distribution, but does not provide an account of how exactly the drawing process takes place. It goes beyond a statement that the world we live in corresponds to one of the attainable states, as it extends to the idea that all of of attainable states are equally likely. The epistemic view largely treats processual concerns as dispensable, although (as we discuss below in Sect. 5) they may also be argued in some instance to be captured by the description of the constraints. This is a high price to pay if one is interested in explanations, since these depend critically on an account of processes[10].

In principle, the processual and the epistemic accounts are compatible, but they involve very different underpinnings. Indeed, the epistemic account has been defended as an alternative to the processual account [28]. Our larger case in this paper is that adequate explanation requires both attention to processes and an adequate specification of them. Statistical physics can contribute to explanation the social sciences only if it attends to both.

[9] See e.g. Banerjee and Yakovenko [1].

[10] It is important to note that it is possible to be a Bayesian and yet hold that an account of processes is necessary for *explanation*, as we do here. Alternative accounts of processes can themselves be the subject of alternative prior probability assessments. We are therefore not simply making a "frequentist" critique.

4 Challenges in the application of statistical mechanical methods in the social sciences

Both the epistemic and the processual views involve implicit assumptions. Although these assumptions are potentially problematic even in specific natural science contexts, they may be especially so in the context of human economic and social life. In arguing this, we echo the points of view of Gallegati et al. [20] and of Ormerod [39] which argue for taking note of the specificities of economic life (for instance, that it involves production and not merely exchange) in order to have more realistic applications of statistical physics to economics. We also align with them in arguing too for taking note of prior work by economists – and one might add, in the social sciences more generally – which can inform the sensible interpretation and development of such applications. To be clear, we are not arguing that the issues we point to are relevant only for the application of statistical physics approaches to economics, but they certainly are relevant to them.

4.1 Agency

The most basic reason for difficulty in applying metaphors from the physical sciences straightforwardly to the social sciences is what may be called the fact of human agency, which has both individual and collective aspects. Human action is intentional (or effectively experienced as such). It is as a consequence directional. Human action cannot be likened to a motion of particles bringing about entirely random collisions: there are, on the contrary, reasons that particular states are more likely than others. This is reflected in both their individual and collective behaviors. For example, if individual firms in a capitalist system are thought of as pursuing profits, then even if they do not do this wholly efficaciously, the presence of such a motive will generate directional consequences, such as a propensity for them to pursue strategies that lower costs and raise revenue. This may in turn give rise to phenomena such as arbitrage (e.g. a tendency to shift capital from industries where profit rates are lower to those where they are higher, or to buy goods where they have a lower price and to sell them where they have a higher price). Further, such firms acting individually or jointly may be expected to favor, and to seek, arrangements that increase their collective profitability (e.g. political lobbying to lower tax rates). Although some such behaviours may be possible to accommodate within a statistical mechanics formulation (see e.g. Shaikh and Foley [15], who do exactly this, respectively through drift-diffusion dynamics and entropy maximization model) the question of how to understand the motivation of actors, which if they were to vary might lead to different systemic outcomes, remains crucial to explanation. This is therefore not an argument against a statistical characterization of behaviour but is rather an argument for a suitable characterization of it, based upon a contextually appropriate understanding.

What is the source of such intentional or motivated behavior? Social scientists have given much attention to this issue, and offered varying interpretations, with the role assigned to systemic factors and formative contexts in shaping individual action, and how these are understood to have such effects, varying across theories. For example, mainstream economists have given much importance to the role of individual "preferences" in guiding choices, but paid little attention to how they come about. In contrast sociologists have given very considerable attention to the manner in which individual's action-guiding motivations, attitudes, inclinations, and beliefs may be formed in a social context[11]. Other disciplines have emphasized the idea that

[11] For instance, the idea of a "habitus" conditioning an individual's worldview and choices (Bourdieu, [4]) or the idea of a mode of production which shapes the ruling ideas of an age [36].

some drives are biologically rooted[12]. Although there may not be consensus either about what motivations prevail in a given setting or about their origin, there is a general recognition that they matter in social life, and cannot be altogether ignored, but this in turn raises questions of how to distinguish contexts from one another, and of the judgments to be employed in doing so. The description of the dynamics of a system may be shaped by the description of the system (for instance, "Capitalism", "Japanese Capitalism", or "Japanese Capitalism in the 1970s"). Ultimately, the choice of models and their precise characterization depends on understandings of what the relevant goals, drives etc. are, and these may well also depend crucially on contextually relevant judgments.

4.2 Changeable "constraints"

The formative contexts that inform individual action – e.g. culture, social structures, institutions, political economy – might be thought to generate the features (the "constraints") that define the system and its attainable states, but also to shape the behaviors of agents who operate within these constraints (at least those that are present at any one time). In turn the behaviors of those agents may shape these features too, whether through gradual influence over historical time or through more immediate "revolutionary" transformation, or other intermediate possibilities. From this perspective, the apparent constraints may not in fact be truly constraining. The behavior of the "molecules" within a "box" and the properties of the "box" (in particular forces operating upon it to change it) may have to be taken note of together. The point is not simply that due to changes of regime, there are occasional switches which must be taken into account (as suggested for instance by Gallegati [20]), for which there may well be suitable technical tools. It is rather that regimes are continually in the process of endogenous transformation.

The possibility of such mutual causal influence of societal "structure" and individual "agency" has led some social scientists to suggest that the distinction should be collapsed (see e.g. Giddens [23], [56]). If structure influences agency and agency influences structure then the "rules" that describe how the system operates, and that govern change, may themselves be changeable under these mutual influences[13]. The "data generating process" giving rise to observed economic and social outcomes may therefore lack stationarity, let alone ergodicity, over historical time[14]. While the application of the "maximum entropy" methodology does not require an assumption of ergodicity, as underlined by Jaynes, it does require the identification of the specific constraints that apply at a moment in time. Although this is no embarrassment, but simply in the nature of the exercise, which are the relevant constraints is a matter of judgment. It involves for instance the identification of those facts about the world that are slowly changing, or difficult to change, and which therefore may be treated for as fixed parameters which are to be respected and which may be treated as changeable. A revolution, or a reform, may cause the assumed "constraints" to melt away.

What features of the world are viewed as "constraints" will depend on the empirical situation but also on the role of the analyst (see e.g. [41]).

5 Processual explanations?

Providing a processual interpretation of the applicability of statistical mechanics to economics requires identifying processes that can be expected to give rise over time to

[12] Ideas associated, for example, with Skinner [55], Wilson [60], or Freud [18].

[13] Unger and Smolin [57] apply the idea that "change changes" to the universe itself.

[14] On the non-ergodic nature of economic life, see e.g. Davidson [5].

the specific anticipated outcomes. For instance, the assumption that molecules randomly collide and give rise to consequent transfers of energy and momentum, often motivates the classical statistical mechanical theory, and underpinned results such as the "H Theorem" which played an important historical role in the justification of the Second Law of Thermodynamics. The theorem demonstrates that if probabilities of transition between accessible states are symmetric then the likelihood of the system being in any one of the states will tend toward equalization (maximization of entropy)[15]. The theorem connects a process (symmetric transition probabilities) with an outcome (the probability distribution of accessible states defined by maximization of entropy) and thereby provides the elements of an explanation. Of course, an explanation can generally be deepened, for instance in this case by explaining why transitions between accessible states can be treated as being probabilistic and symmetric.

It is far from obvious, however, that such results of statistical mechanics, even if deemed applicable to the physical world, can be imported to provide justification for the application of statistical mechanical methods to the economic and social world. For instance, if we were to ask whether analogous assumptions to those used to prove the "H Theorem" hold in the social world, we would have to take note that in the economic and social world the transition probabilities between accessible states are not symmetric. There may be systematic individual motivations, social orientations and institutional tendencies (e.g. for firms to seek higher profits, for lobbies to promote those interests, for governments to respond to those efforts, for workers or organize on the basis of perceived commonalities, etc.) all of which make the likelihood of transition between elements in any pair of accessible states greater in one direction than another and this may indeed be necessary to maintain an unequal societal situation. For instance, the symmetry hypothesis would appear to imply that the chance of a specific poor black man exchanging social and economic ranks with a specific rich white man (the scenario considered in the popular film Trading Places) is the same as the chance of their switching back[16]. However, the chance of a transition occurring in the first direction is likely to be greater than the chance of its occurring in the second direction, if social dynamics relating to race or class make transitions downward less likely than transitions upward for rich whites, and making transitions downward more likely than transitions upward for poor blacks. Similar factors may influence other "systemic" features (for instance, policy reforms influencing taxes and regulation; it may be more likely in a given political environment that taxes are revised downward than that they are revised upward even though these may both be "accessible states"). Recognizing the presence of such directional factors is often integral to understanding social processes, and therefore to providing explanation of social outcomes, just as the random motion of particles plays a role in providing processual explanations of statistical mechanics. The non-symmetry of transition probabilities between accessible states is a reflection of the existence of human agency and the tendency for it to be exercised in specific ways (shaped in turn by contexts).

Directional behaviors may help to bring about an equilibrium, but may also undermine any movement toward an equilibrium. Observed outcomes may as a result not be equilibria, nor even be entropy increasing, let alone maximizing, within a specific setting (e.g. within the boundaries of a single country). The second law of thermodynamics ensures that entropy must be globally increasing, but this is compatible

---

[15] See the proof in Reif [42], Appendix A.12.

[16] Arguably, it is because symmetric probabilities of transition between accessible states are so improbable in the economic and social world that their occurrence provides a rich basis for speculative fiction and film. Mark Twain's The Prince and the Pauper provides another famous example of such a scenario.

with a local increase. For instance, economic growth and development of countries can generate increasingly complex and productive forms of economic organization, but may do so at the cost of using up concentrated energy sources and generating waste. This is a special case of the broader example of life itself as a form of local entropy decrease[17]. Intentions and drives may be relevant to bringing about such local entropy decreases.

The epistemic justification takes entropy maximization to be the appropriate methodological stance of an analyst given that all available knowledge is fully taken account of in the specified constraints. As noted earlier, this distribution is privileged because it corresponds to the assignment of equal likelihoods to all attainable states, and therefore to an "equiprobability" assumption. There is no demand within this approach for an account of the processes operating within the system being studied. The absence of any processual account, however, undermines the ability of the approach to provide an explanation. Elements of an explanation may still be attained through a specification of the constraints, suggesting a specific processual interpretation, even if no such account is directly provided, but such an account is likely to remain incomplete. For example, in an exercise described in Foley and Scharfenaker [16], entropy maximization is motivated by the idea that a given industry-level profit rate is attained while permitting variations in that attained by individual firms. This allows for the interpretation that firms are seeking to further profits but succeeding in doing so to different degrees, with their dispersion determined by the requirement that in any attainable state the average profit rate constraint must be satisfied. The departures from a uniform profit rate that results (unlike in the widely accepted "neoclassical" model) can be given various interpretations, for instance that they arise due to the propensity of firms to commit errors, thus departing from strict "maximization". Although such a perspective may seem to offer insight as to why a profit distribution rather than a single profit rate prevails, much still depends on the framework imposed by the analyst, calling for judgment about what is socially and economically salient in the given context. There is a duality between the representation of a distribution as resulting from agents maximizing a payoff (e.g. profits) subject to entropy being at least a minimum level, and the maximization of entropy subject to the agents' payoffs being at least a minimum level. In the first instance, the minimum entropy constraint may be viewed as arising due to a likelihood of the agents making mistakes or otherwise deviating from maximization. In the latter instance, the minimum payoff constraint may be viewed as corresponding to a participation constraint on the part of firms. Although these two representations are identical in their distributional consequences, they entail very different understandings of the *reasons* why a particular pattern emerges. In other words, although the distribution is pinned down by the formalism, the explanation as to why it holds is not, if by an explanation we mean to refer to processes and causes. It is clear that the maximum entropy formalism is helpful here, but it remains incomplete from this standpoint.

[17] "…the second law of thermodynamics goes in the wrong direction; if the second law were the driving principle, evolution would proceed inexorably back to the primordial soup, which has a much higher entropy than would any collection of living creatures that might be made from the same atoms." [29]. Of course, much depends on how the system is specified by the analyst, which underlines the roles of judgment and purpose in the exercise: "I have been asked several times whether, in my opinion, a biological system, say a cat, which converts inanimate food into a highly organized structure and behavior, represents a violation of the second law. The answer I always give is that, until we specify the set of parameters which define the *thermodynamic state* of the cat, no definite question has been asked!" [27].

6 History vs. equilibrium

The potential difficulty of applying the concept of equilibrium as it arises in statistical mechanics to economics can be seen from various perspectives, one of which involves a famous contrast between history and equilibrium [25,43]. The use of equilibrium reasoning in relation to entropy maximization in statistical mechanics is sometimes motivated by Liouville's Theorem, which states that if all accessible states are equally probable that will remain the case over time (see e.g. [42]). This is the result of the more general principle that if the distribution of the system over accessible states is a function of some time-invariant constant of the system (e.g. energy) then it too will remain constant over time.

Although such an idea may seem general, in the economic and social world there may be no such readily identifiable systemic invariants (conserved quantities corresponding to a "Hamiltonian"). The fundamental reason is the existence of human agency. Even though human beings in society are subject to physical laws, they can act in ways, individually and collectively, that upset any seeming invariant of social and economic life. "Social science" has yet to identify such invariants over any duration. Even if all accessible states are equally probable to begin with, the exercise of individual and collective agency can lead to "spontaneous" symmetry breaking. Intentional interactions of persons can lead to some states becoming more likely than others, disrupting equilibrium and creating history, even if has seemed to have "ended" (see e.g. the discussion of the "last man" in Fukuyama [19]). "Equilibrium" may not be attained, or it may not be sustained. From this perspective, disequilibrium may, at least in certain contexts and moments, be more pertinent than equilibrium. Although the concept of equilibrium has enjoyed considerable influence in the social sciences, and in particular in economics, it has also been subject to question. A long tradition has attributed special explanatory significance to the concept of equilibrium (e.g. partial or general equilibrium theories of supply and demand, strategic equilibria of game theory etc.). On the other hand, an equally long tradition has emphasized the importance of disruptive phenomena emerging from outside a system and from within it (e.g. unruly dynamics caused by the accumulation processes internal to an economic system).

Even within a perspective in which equilibrium is central, path-dependent processes can be of considerable significance in economic and social life, making the particular equilibria which arise a consequence of history. Both the specific constraints which hold and the processes which are at play may be shaped by it. One example which helps to illustrate the point is that of the agglomeration economy used to understand economic geography, e.g. the origins and economic specialization of cities (see e.g. [32,35]). In an agglomeration economy, the positive spillovers of specific activities can give rise to an incentive to locate close to others, for example because they increase productivity or demand experienced by firms. If only weak interactions between particles are assumed this can be enough to generate random transitions but insufficient to generate concentrations. An agglomeration economy is in contrast an instance in which the forces generating concentration are quite important.

In the presence of such phenomena, path dependence matters, for example determining location and nature of dominant industries (Krugman, op cit). This is perhaps unsurprising. Physical and biological phenomena (e.g. the formation of crystals, or evolutionary dynamics leading to specific life forms emerging) also require "historical" analysis in order to be understood in their specifics, even if some abstract reasoning can be applied across contexts and cases. This is one reason that many economic and social processes seem to be best thought of as "non-ergodic" processes, which have no tendency to converge asymptotically to a specific determinable outcome (an "ensemble average") but which rather can evolve in different directions depending on what is experienced [5].

7 A case: the colonization of the Americas

An interesting case-study in the application of statistical mechanical reasoning which shows the role of such reasoning but also the tension in which it stands to historical analysis is offered by the "discovery" and colonization of the Americas. Although this is admittedly not a case which the literature on statistical physics as applied to economics has addressed, it illustrates the issues well.

According to a possible interpretation using a conventional "neoclassical" framework, the "discovery" of the Americas and its subsequent interaction with the old world meant that a labor abundant and land scarce region met a labor scarce and land abundant region (for relevant insights, see e.g. [13,14,59]). The consequence was movement of persons (voluntary and forced) from one region to another. New opportunities for trade and flows of factors of production (in particular, labour) were opened by differences between regions in factor proportions, arising from their different natural and social endowments. The resulting movement of goods and factors could be seen as an "entropy increasing" process (much like one caused by removing a divider separating a box that contains a gas from one that does not) which makes it possible for previously inaccessible states to be realized. The result of such a process, as it tends toward entropy maximization, is that states which arise from flows having taken place from one side of the "divider" to the other become vastly more likely, with the states which are most likely to be realized being those which involve the gas molecules being nearly evenly distributed across the space of the newly enlarged box. The economic counterpart of such a process, envisioned in the "standard" literature on trade and factor flows, is a new free trade equilibrium or, in the presence of factor flows, an "integrated world equilibrium" (see [7]). A sharp probability distribution might be expected around the integrated equilibrium, since states approaching autarky are highly improbable to be sustained given either intentional or random interactions and transitions. It is interesting to note that in economics, as in statistical mechanics, such a process can be irreversible (i.e. restoring the divider needs not restore the status-quo ante) in particular in the case in which integration provides for the flow of factors as well as goods (consider for instance the flows of people, flora, and fauna between the old and the new worlds).

In real time, however, as distinguished from instantaneous "model time" it may take centuries to get to such a point, and it may not be arrived at even then (e.g. the population density of the Americas continues to be rather lower than for Europe). Perhaps this is both because there remain impediments to such flows, and because there are "dis-equalizing" forces operating alongside equalizing ones. However, not all of what is observed corresponds simply with "equalization". In order to understand the features that are observed we must pay attention to history. For example, North-Eastern Brazil, may have become a center of relative poverty within the country as a historical consequence of the end of a prior boom in the production of sugar on plantation agriculture in the eighteenth and nineteenth century, which had led to the large-scale importation of slaves and the development of a substantial population in these areas (see e.g. [17]). Meanwhile, a part of the new world (the United States) became richer than its old-world mother country (the UK) overtaking it and Europe in general to become a new leading country and an economic center of gravity for the world economy. It was able to do so as a direct result of its access to a giant untapped and continental-sized resource based which it could combine with new institutional arrangements. Due to these "advantages of backwardness" [22] a peripheral country eventually became central. Politically enforced labor immobility has meanwhile diminished equalization (e.g. controls on movement of people from Latin America and Europe to North America). These facts cannot be straightforwardly interpreted on the basis of "entropy maximization". While the statistical mechanical lens appears to provide a possible lens for a very gross interpretation of the consequences of the removal of the "divider" it does not adequately explain the phenomena which are of ultimate interest. This is not merely a matter of getting the "second order" details right but of understanding some "first order" phenomena which would otherwise be neglected, such as the emergence of new centers of growth (North America) giving rise to disequalization and "overshooting" of equilibrium. "Equilibrium" can only provide partial and provisional insight, and must be accompanied by and understanding of the diverse economic and social factors that are at play in "History".

8 Finding what you look for?

One argument that has been presented by proponents of statistical mechanical approaches is that the results generated by "entropy maximization" frameworks fit the actual data extraordinarily well. They make a claim, accordingly, of unreasonable effectiveness of the method, meant to justify its use (see Gallegati et al. [20], [1]). A sceptic may, however, point to the possibility that the techniques used constitute in effect a sophisticated retrospective curve-fitting method. Its ability to "fit the data" is from this perspective no demonstration of explanatory efficacy. For a very broad class of distributions, identification of a sufficient number of moment constraints can suffice to estimate a distribution with some accuracy [3,30,37]. There is a literature on the use of entropy maximization as a non-parametric estimation technique in which the constraints used provide relevant input from the data. The Kullback-Leibler divergence, widely used as a measure of goodness of fit in the statistical physics and economics literature, provides a criterion for how well a model fits the data. The high goodness of fit can be understood as deriving in part from the fact that the constraints convey much information relevant for recovering a distribution – such as its moments, which are derived from the data itself[18].

The fact that a maximum entropy distribution corresponding to specific assumed constraints appears to fit the data quite well may be useful in providing a way to summarize the data (although its usefulness for this purpose should be compared with that of other available techniques for data description). But the ability to provide a seemingly "good fit" to the data ex post is not in itself a justification of the maximum entropy technique as a mode of economic explanation. How then should the validity of the theory be tested? It might be better to undertake a direct comparison with deliberately considered alternative models. In this, we are in agreement with Gallegati et al. [20] who underline the need for an "explicit test of a well-defined hypothesis" and who are similarly critical of "self-deception by a combination of data-mining and apparently good visual fits". The literature has generally not provided such tests or comparisons, instead relying on visual inspection of the relationship between actual data and the distributions fit to that same data on the basis of statistical physics models. Such visual inspection procedures are often, problematically, also applied only to parts of the distribution (Ormerod [39]). Such procedures may be predisposed to theory confirmation. In any case, goodness of fit does not in itself suffice for explanation, which requires attention to causes and processes.

[18] Indeed, the model parameters that minimize the KL divergence between the model and the observed data also maximize the log-likelihood of the observed data, placing the technique in a familiar context of statistical inference.

9 Conclusions

Statistical mechanics reasoning may be applicable in the economic and social sciences, but only if adequate consideration is paid to the specific contexts and conditions of its application. This requires attention to "non-mechanical" processes of interaction, inflected by power, culture, institutions etc., and therefore of specific histories which gives rise to these factors. It is unlikely that a unified account will be found which possesses explanatory power across all of time and space. Perhaps no one has made this claim, but it is nevertheless important to consider what is the proper domain of application of a theory.

Our more general contention is that, outside of very specific cases, statistical physics is more likely to provide useful metaphors and ways of thinking than computational techniques and definite answers. Although statistical mechanical explanation can shed light on particular mechanisms that may be at play (e.g. explaining how distributions are shaped by the differences between the ways in which wages and profits may respectively evolve) these must be interpreted in a specific causal context. For instance, capitalism as a system gives rise to the wage-profit distinction, which does not exist in a comparable way in petty-commodity production or in feudalism). Both in the natural sciences and the social sciences, "reductionist" explanations (for instance, which interpret a system in terms of the interactions of molecules or individuals) apply in the context of a specific set of conditions (often called "laws" in the natural sciences) that define the behaviors of the parts. These may in turn be associated with a specific "system" (we therefore speak of the behavior of molecules in a "gas" or of firms in "capitalism"). The "macroscopic" properties of a system and the "microscopic" workings of its elements cannot therefore be separated. Each such system may also be a particular rather than a universal phenomenon. In the context of the human sciences, agency – not merely the exercise of individual choice but the shaping of the circumstances of collective life – is the central factor both in determining the properties of a system and in shaping individual choices. The exercise of human agency may bring a social system closer to an "equilibrium" in some circumstances, and disrupt it in others. It is among the specifically social factors that must be taken note of in the dialogue between statistical physics and the social sciences.

I am grateful to two anonymous referees for very helpful comments. I would like to thank Ellis Scharfenaker and Jangho Yang for encouragement to complete the draft and Ravi Kanbur, Paulo dos Santos, Duncan Foley, Anwar Shaikh and Venkat Venkatasubramanian, among others, for clarifying discussions. Valuable research assistance was provided by Mohamed Obaidy and Xingxing Yang. This research was independent, and was not supported by any research grant.

References

1. A. Banerjee, V. Yakovenko. New J. Phys. **12**, 075032 (2010)
2. G.S. Becker, J. Political Econ. **70**, 1 (1962)
3. P. Biswas, A.K. Bhattacharya, J. Phys. A: Math. Theor. **43**, 405003 (2010)
4. P. Bourdieu, *Outline of a Theory of Practice* (Cambridge University Press, Cambridge, 1977)
5. P. Davidson, Uncertainty in economics, in *Encyclopedia of Statistical Sciences*, edited by S. Kotz, C.B. Read, N. Balakrishnan, B. Vidakovic, N.L. Johnson (Wiley, 2012)

6. A. Deaton, N. Cartwright, Soc. Sci. Med. **210**, 2 (2018)
7. A. Dixit, V. Norman, *Theory of International Trade* (Cambridge University Press, Cambridge, 1980)
8. P. dos Santos, Complexity **2017**, 8358909 (2017)
9. P. dos Santos, Eur. Phys. J. Special Topics **229**, 1603 (2020)
10. P. dos Santos, N. Weiner, Entropy **21**, 367 (2019)
11. A. Dragulescu and Y. Yakovenko, Statistical mechanics of money, income, and wealth: A short survey, in *Modeling of Complex Systems: Seventh Granada Lectures, AIP Conference Proceedings* (New York, 2003), Vol. 661, pp. 180–183
12. N. Elias, *The Society of Individuals* (Basil Blackwell, Oxford, 1991)
13. R. Findlay, The Triangular trade and the Atlantic economy of the eighteenth century: A simple general-equilibrium model, in *Essays in International Finance, No. 177* (Princeton University, 1990)
14. R. Findlay, K. O'Rourke, *Power and Plenty: Trade, War, and the World Economy in the Second Millennium* (Princeton University Press, Princeton, NJ, 2007)
15. D. Foley, Eur. Phys. J. Special Topics **229**, 1591 (2020)
16. D. Foley, E. Scharfenaker, Maximum entropy estimation of statistical equilibrium in economic quantal response models, Working Paper 1710, New School for Social Research, Department of Economics, 2017
17. A.G. Frank, *Latin America: Underdevelopment or Revolution* (Monthly Review Press, New York, 1969)
18. S. Freud, *Civilization and its Discontents*, translated by J. Strachey (W.W. Norton, New York, 1961)
19. F. Fukuyama, *The End of History and the Last Man* (The Free Press, New York, 1992)
20. M. Gallegati et al., Physica A **370**, 1 (2006)
21. N. Georgescu-Roegen, *The Entropy Law and the Economic Process* (Harvard University Press, Cambridge, MA, 1971)
22. A. Gerschenkron, *Economic Backwardness in Historical Perspective* (Belknap Press of Harvard University Press, Cambridge, MA, 1962)
23. A. Giddens, *The constitution of society: Outline of the theory of structuration* (Polity Press, Cambridge, 1984)
24. S.J. Gould, *The Structure of Evolutionary Theory* (Harvard University Press, Cambridge, 2002)
25. D. Harris, Joan Robinson on "History versus Equilibrium", in *Joan Robinson's Economics A Centennial Celebration*, edited by B. Gibson (Edward Elgar, 2005)
26. C.G. Hempel, P. Oppenheim, Philos. Sci. **15**, 2 (1948)
27. E.T. Jaynes, Gibbs vs. Boltzmann Entropies (1964), in *Papers on Probability, Statistics, and Statistical Physics*, edited by R.D. Rosenkrantz (D. Reidel, Dordrecht, 1983)
28. E.T. Jaynes, Foundations of probability theory and statistical mechanics, The Delaware Lecture (1967), in *Papers on Probability, Statistics, and Statistical Physics*, edited by R.D. Rosenkrantz (D. Reidel, Dordrecht, 1983)
29. E.T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003)
30. V. John, Chem. Eng. Sci. **62**, 11 (2007)
31. J.M. Keynes, *Treatise on Probability* (Macmillan & Co., London, 1921)
32. P. Krugman, J. Political Econ. **99**, 483 (1991)
33. S. Lukes, Br. J. Sociol. **19**, 2 (1968)
34. E. Malinvaud, *Theory of Unemployment Reconsidered* (John Wiley and Sons, New York, 1977)
35. A. Marshall, *Principles of Economics*, revised edn. (Macmillan, London, 1920)
36. K. Marx, F. Engels, *The German Ideology* (Prometheus Books, New York, 1846, 2011)
37. J.L. Muñoz-Cobo, R. Mendizábal, A. Miquel, C. Berna, A. Escrivá, Entropy **19**, 486 (2017)
38. W.J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, B. Yu, PNAS **116**, 22071 (2019)
39. P. Ormerod. Eur. Phys. J. Special Topics **225**, 3281 (2016)
40. G. Poitras, Physica A **507**, 89 (2018)

41. S. Reddy, J. Ethics **9**, 119 (2005)
42. F. Reif, *Fundamentals of Statistical and Thermal Physics* (Mc-Graw Hill, New York, 1965)
43. J. Robinson, History vs. Equilibrium, in *Contributions to Modern Economics* (Blackwell, Oxford, 1978)
44. J.B. Rosser, Eur. Phys. J. Special Topics **225**, 17 (2016)
45. P.A. Samuelson, *Foundations of Economic Analysis* (Harvard University Press, Cambridge, MA, 1947)
46. G. Semieniuk, E. Scharfenaker, Metroeconomica **68**, 465 (2017)
47. A. Sen, Oxford Econ. Pap. **32**, 353 (1980)
48. A. Sen, Philos. Trans. R. Soc. London **407**, 3 (1986)
49. A. Sen, Soc. Res. **71**, 583 (2004)
50. A. Shaikh, A. Ragab, An international comparison of the incomes of the vast majority, Working Paper, The New School, 2011
51. A. Shaikh, N. Papanikolau, N. Weiner, Physica A **415**, 54 (2014)
52. A. Shaikh, Rev. Political Econ. **29**, 1 (2017)
53. A. Shaikh, Eur. Phys. J. Special Topics **229**, 1675 (2020)
54. A. Shaikh, Rethinking microeconomics: A proposed reconstruction, The New School Economics Department Working Paper 06/2012, 2012
55. B.F. Skinner, *Beyond Freedom and Dignity* (Knopf, New York, 1971)
56. R.M. Unger, *Social Theory: Its Situation and Its Task, Volume 2 of Politics: A Work in Constructive Social Theory* (Cambridge University Press, Cambridge, 1987)
57. R.M. Unger and L. Smolin, *The Singular Universe and the Reality of Time: A Proposal in Natural Philosophy* (Cambridge University Press, Cambridge, 2014)
58. V. Venkatasubramanian, *How Much Inequality is Fair* (Columbia University Press, New York, 2017)
59. J.G. Williamson, J. Econ. History **62**, 55 (2002)
60. E.O. Wilson, *Sociobiology: A New Synthesis*, 2nd edn. (Harvard University Press, Cambridge, MA, 2000)
61. J. Woodward, Scientific explanation, in *Stanford Encyclopedia of Philosophy*, edited by E.N. Zalta (2017), https://plato.stanford.edu/archives/fall2017/entries/scientific-explanation/