How-Possibly Explanations in Economics: Anything Goes?∗

Till Grüne-Yanoff† Philippe Verreault-Julien‡

20th November 2020

Abstract

The recent literature on economic models has rejected the traditional requirement that their epistemic value necessarily depended on them offering actual explanations of phenomena. Contributors to that literature have argued that many models do not aim at providing how-actually explanations, but instead how-possibly explanations. However, how to assess the epistemic value of HPEs remains an open question. We present a programmatic approach to answering it. We first introduce a conceptual framework that distinguishes how-actually explanations from how-possibly explanations and that further differentiates between epistemic and objective how-possibly explanations. Secondly, we show how that framework can be used for methodological appraisal as well as for understanding methodological controversies.

Contents

1 Introduction 2
2 Anything Goes? 2
3 Epistemically Possible HPEs: Krugman on Trade and Geography 5
4 Objectively Possible HPEs: Axelrod et al. on Social Networks and Cooperation 7
5 Conclusion 9
References 10

∗Penultimate version. Final version to appear in the Journal of Economic Methodology. The authors are listed in alphabetical order. We would like to thank Ylwa Sjölin Wirling for helpful feedback on a previous version of the manuscript. Till Grüne-Yanoff acknowledges support from the Swedish Research Council (Vetenskapsrådet, contract number 2018-01353). Philippe Verreault-Julien acknowledges support from the Fonds de recherche du Québec – Société et culture (FRQSC).
†Royal Institute of Technology (KTH) Stockholm. Email: gryne@kth.se. Website: http://people.kth.se/~gryne/.
‡London School of Economics and Political Science. Email: p.verreault-julien@lse.ac.uk. Website: http://pvjulien.net.
1 Introduction

What makes a theoretical model in economics good? Many methodologists held that models have to provide explanations of actual phenomena for them to have any epistemic value. Since we have good reasons to believe that many models neither explain (Reiss 2012, 2013) nor provide predictions of actual phenomena, this led some to the rather pessimistic assessment that many economic modelling practices have no epistemic value. At best, they can be appraised for their “heuristic”, pedagogical or otherwise non-epistemic uses (Alexandrova 2008; Alexandrova and Northcott 2013; Fumagalli 2016; Hausman 1992).

A growing number of authors have instead rejected the assumption that having an explanation of actual phenomena is necessary for epistemic value. Instead, they have argued that many scientific modelling efforts aim not at how-actually explanations (HAEs) but at how-possibly explanations (HPEs) (Aydinonat 2007, 2018; Aydinonat and Köksal 2019; Basso, Lisciandra, and Marchionni 2017; Grüne-Yanoff 2013a, 2013b; Kuorikoski and Ylikoski 2015; Mäki 2009; Reiss 2008, ch. 6; Rohwer and Rice 2013; Verreault-Julien 2017; Weisberg 2013, ch. 7; Ylikoski and Aydinonat 2014).

However, how to assess the epistemic value of HPEs is still by and large an open question. In this paper, we present a programmatic approach to answering it. We first introduce a conceptual framework based on our previous work (e.g. Grüne-Yanoff 2009; Sjölin Wirling and Grüne-Yanoff forthcoming; Verreault-Julien 2019) that demarcates HPEs from HAEs, but also distinguishes two subspecies of HPEs with distinct epistemic value. Illustrating with various cases from economics, we secondly show how that framework has the resources to guide our methodological appraisal of economic modelling and to illuminate methodological controversies.

2 Anything Goes?

The descriptive claim that economics often provides HPEs and not HAEs raises one conspicuous methodological worry: Why should we believe that HPEs are epistemically valuable in the first place? It may well be true that economists often only intend to provide HPEs, but that does not make it a sound scientific practice. We believe this concern is by and large misguided. This is because HPEs have been considered valuable for several reasons in the literature. For instance, they...

- can show what could have produced a known phenomenon (Brandon 1990, ch. 5).
- can be heuristically useful because they tell where to look for empirical evidence (Resnik 1991).
- test impossibility or necessity claims (Grüne-Yanoff 2009; see also Reiss 2008, ch. 6).
- are a bona fide type of explanation (Dray 1968; Forber 2010).
- add to a ‘menu’ of explanations (Aydinonat 2018; Aydinonat and Köksal 2019; Ylikoski and Aydinonat 2014).
- serve for inference to the best explanation (Lipton 2004).
• may provide understanding (e.g. Bokulich 2016; Kuorikoski and Ylikoski 2015; Rice 2016; Verreault-Julien 2017, 2019; Ylikoski and Aydinonat 2014).

That said, even though we may have good reasons to hold that HPEs can be valuable, one could still wonder whether we are letting economists off the hook by claiming that ‘mere’ HPEs are epistemically valuable. Are we not giving economists a license to come up with any fanciful story they like? These are legitimate worries, but worries we consider can be addressed. This is because, methodologically, it is not the case that anything goes; there are conceptual resources at our disposal to assess the epistemic contribution of HPEs.

HPEs are possible explanations (Verreault-Julien 2019). HPEs often show what is the possible cause of a phenomenon. This suggests a first criterion to assess HPEs: good HPEs are true HPEs. HPEs, like any modal claim, are true or false. For example, ‘it is possible that General Motors discontinued its electric car EV in 2002 due to a conspiracy of the oil industry’ is either true or false. Some HPEs are thus akin to just-so stories that express a false possibility claim. Other HPEs state a true possibility claim. Truth and falsehood are one primary coarse-grained way of appraising HPEs. Yet this point has often been overlooked in the literature (cf. Verreault-Julien 2019). By emphasizing that HPEs need to be true already goes a long way in responding to the worry that any fanciful story may count as an HPE.

If all false HPEs are bad, true HPEs are not all equally good. Recall that HPEs are possible explanations. A key implication is that different possibility claims may have different truth conditions. This is because the possibility operator (♦) can have multiple interpretations (Kment 2017). Typically, economists are interested in what is economically possible (i.e. roughly, what might result from the economic causes operational in a given system). But possibility claims can also be, for instance, merely logically or conceptually possible. So, the same claim can be true under a given interpretation of the modal operator and false under another. For example, we may grant that it is logically possible for GM to have discontinued its electric car EV in 2002 due to pressure from Opus Dei. But we may deny its economic possibility: in 2002, Opus Dei did not have enough economic influence to stop a major company from pursuing a project that was economically viable. Claims of logical possibility are usually not very informative because logical possibility does not constrain much. More restrictive modalities, like physical or economic possibility, instead tell us what is in accord with the broadly construed physical or economic causes and not merely our logic. Thus, when evaluating HPEs it is very important to be clear and explicit about the interpretation of the modal operator. This is another feature of HPEs that has been neglected in the literature (Sjölin Wirling and Grüne-Yanoff forthcoming; Verreault-Julien 2019).

1 In the following we will bracket issues pertaining to the epistemology of modality. How we come to know modal claims or what is the source of their justification are thorny philosophical issues which are of course beyond the scope of this paper. Proposed but controversial accounts include (i) satisfaction of nomic constraints (Fischer 2017; Massimi 2019) and (ii) appropriately constrained versions of imaginability (Grüne-Yanoff 2009; Kung 2010; Williamson 2007). What is clear, nevertheless, is that 1) people often make modal claims and 2) these claims can be true or false: e.g., that ‘it is possible to break that coffee mug’ has a truth value.

2 See Hubálek (2020) for an overview of the different meanings of ‘just-so stories’.
There is yet another crucial distinction related to the interpretation of the modal operator and that concerns the truth conditions of HPEs. Some HPEs are *epistemically* possible whereas others are *objectively* possible (Sjölin Wirling and Grüne-Yanoff forthcoming). Epistemic possibility depends on knowledge of the actual state of the world. A proposition is epistemically possible if it is not ruled out by an agent’s (or epistemic community) knowledge about what is actually the case. Establishing epistemically possible HPEs (EpHPEs) typically constitutes epistemic progress on the way towards HAES. We are often interested in knowing ‘Why *p*?’; for instance ‘Why does the tiger bush display a stripe pattern?’ (Bokulich 2014). If we do not know what the explanans is, we might contrive a potential candidate *c* (e.g. a factor that implies the difference the explanatory question is about) and then use the knowledge we have in order to determine whether such a candidate is indeed epistemically possible. If, for example, we know that *c* does not exist in the tiger bush environment, or that its effect on the stripe pattern is inhibited by something else, then we must conclude that *c* is *not* a possible explanation of the stripe pattern. If knowledge of the actual rules out the epistemic possibility of the candidate, then we have to look for other candidates. If, however, it is possible, we may search for evidence whether it is actually true. Establishing that an EpHPE is not only possible, but also actual, yields an HAE of the phenomenon, rendering alternative EpHPEs obsolete (Bokulich 2014).

Sometimes we don’t ask ourselves questions like ‘Why *p*?’ but rather questions like ‘Why is *p* not possible?’ and ‘Why is *p* necessary?’. When we ask these questions, EpHPEs won’t do the trick. To answer them, we need more than information about what is possibly actual — instead, we need information about what is possible, although perhaps not actual. Thus, instead of epistemically possible, we need *objectively* possible HPEs (ObHPEs). Contrary to epistemic possibility, objective possibility is to some extent independent of the way things are in the actual world. States of affairs could have been different from what they actually are. Therefore, ObHPEs may be inconsistent with parts of our knowledge of the actual. For example, Grüne-Yanoff (2009) argues that the checkerboard model of residential segregation (Schelling 1971, 1978) allows to refute an impossibility claim by identifying a possible factor — independently of whether this factor is operational in the actual world. What especially matters for assessing ObHPEs is the relevance of the modality for the question at hand. Ways the world could physically or economically be is usually more interesting to scientists than logical or metaphysical possibility. And an ObHPE may be true under one interpretation (e.g. logical possibility) and false under another (e.g. physical possibility).

EpHPEs are particularly important in economics (and science more generally) because we often want our possible explanations to not be ruled out by the evidence. On the one hand, EpHPEs add explanations to the ‘menu’ (Aydinonat 2018; Aydinonat and Köksal 2019; Ylikoski and Aydinonat 2014) and can serve as candidates for inference to the best explanation (Lipton 2004). On the other hand, ObHPEs may inform us about the necessity

---

3Explaining why something is possible or impossible is not the only domain of ObHPEs (Sjölin Wirling and Grüne-Yanoff forthcoming); we focus on it here for ease of exposition.
or contingency of states of affairs (Lipton 2009). In any case, what is crucial to recognize is that once we interpret what is the appropriate notion of ‘possibly’ in a given context, we can provide truth conditions for the possibility claim. In turn, this allows assessing the epistemic contribution of HPEs. In the next two sections, we illustrate the use of this conceptual framework to better understand cases of modelling in economics.

3 Epistemically Possible HPEs: Krugman on Trade and Geography

Until the 1970s, patterns of trade were explained in terms of comparative advantage (Ricardo 1817; see also Ohlin 1933): countries would trade with each other on the basis of their productive differences (e.g. technology). Comparative advantage theory had a considerable explanatory success up until the middle of the 20th century. After that, comparative advantage theory could hardly explain changing trade patterns, namely that countries with similar technology or factor endowments would trade between themselves and within the same industries (intra-industry trade). For instance, countries with similar features would both export and import similar products, e.g. cars. This was incompatible with comparative advantage theory.

Since the prevailing models of trade could not explain the data, a new explanation was needed. One seminal contribution to that new explanation is Krugman (1979, see also 1980). He built a simple model in which increasing economies of scale were the key driver of trade patterns. He summarized his contribution as follows:

> The important point to be gained from this analysis is that economies of scale can be shown to give rise to trade and to gains from trade even when there are no international differences in tastes, technology, or factor endowments. (Krugman 1979, 477)

In 2008, Krugman was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel for his contribution to the development of new trade theory, which provided a novel explanation of patterns of trade based on that model. According to the Prize Committee, Krugman “demonstrated that models based on the assumptions of increasing returns to scale and monopolistic competition can explain important patterns of trade observed in the data” (Royal Swedish Academy of Sciences 2008, 8, our emphasis). Despite the model’s simplicity, the Committee praised its ability to account for the patterns found in the data. It is noteworthy, though, that the Committee acknowledged that empirical tests of the model’s predictive implications “showed mixed results” (Royal Swedish Academy of Sciences 2008, 10).\(^4\) At the time of the award, there was still considerable uncertainty whether the model’s explanans was actually responsible for the trade patterns. Thus, Krugman didn’t obtain the prize for an HAE.

In the terms of our conceptual framework, Krugman’s model provided an EpHPE. It

---

\(^4\)More precisely, of a model integrating Krugman’s account and the traditional factor-proportions theory.
purported to explain actual trade patterns by showing how they could arise even when there is a lack of productive differences between countries. At time of publication, for all we knew, increasing returns and imperfect competition may have explained patterns of trade. The model results were consistent with the data and the empirical evidence did not rule out the possibility that these two factors could produce the patterns of trade. It added another explanation to the ‘menu’ (Aydinonat 2018; Ylikoski and Aydinonat 2014) which could then be tested empirically to determine whether it was also an HAE. In a nutshell, this is a good example of a model that is widely considered to have established a true claim of epistemic possibility.

Not every model that accounts for a phenomenon is a successful EpHPE, however. Such models may fail to provide EpHPEs for either of at least two reasons. First, a model might account for a phenomenon by invoking an explanans that is known not to be actually true. It certainly is possible in some non-epistemic sense that the oil industry put enough pressure on GM to discontinue its electric car EVs in 2002. But if it were known that the oil industry did no such thing, then it could not be a true EpHPE. Perhaps less dramatically, most models that once offered EpHPEs stop doing so when evidence accumulated and it became clear that their explanans were not true. Such is scientific progress.

Secondly, a model may fail to provide an EpHPE because its possible candidate explanans does not imply the difference the explanatory question asks about. That is, although the explanans itself is indeed possible given current knowledge, the explanation is not because the explanans cannot possibly be the difference-making cause of the explanandum. We illustrate this at the hand of Krugman’s (1991) contribution to geographical economics (GE), which has attracted a fair amount (e.g. Dymski 1996; Martin 1999; Sheppard 2001; see also Mäki and Marchionni 2011; Marchionni 2004, 2006, 2012). Krugman developed his core-periphery (CP) model to endogenously explain the location of factors of production. Again, emphasis was put on simplicity and in particular on unification — the ability to account for many seemingly disparate phenomena from e.g. urban and regional economics or traditional economic geography — with the same model. In order to account for diverse phenomena, however, the CP mechanism represented by Krugman’s model had to be stripped down to its bare essentials. Critics seized on this high level of abstraction, arguing that the model failed for purposes of explaining particular agglomerations, because “it is indeed not the unifying mechanism alone that bears the explanatory burden, but the specific ‘details’ too” (Marchionni 2012, 436). In other words, the candidate explanans offered by Krugman’s model, although epistemically possible, falls short of identifying those factors that might produce the difference to be explained (‘why is the agglomeration here and not there?’).

Thus, we can understand the reaction against GE as maintaining that, for all we know, economies of scale and transportation costs alone do not offer an EpHPE of some location phenomena. GE may identify interesting and relevant factors, but reality is more complex and those factors cannot explain the sort of phenomena economic geographers

---

Some of these criticisms are not only directed at Krugman (1991), but more generally at extensions of the CP-model, the actual novelty of GE, and its underlying methodology.
are typically interested in. In other words, the critics were saying that Krugman’s CP model provides a false EpHPE for particular phenomena $p$, in the sense that the claim ‘CP explains $p$’ is not epistemically possible, because the critics claimed to know that other causes contribute to $p$, but are absent in CP.

Again, our goal is not to assess whether the critics’ claim is true. If anything, it seems champions and critics were talking past each other insofar as they were, in fact, seeking to explain different explananda (Marchionni 2006). Indeed, Marchionni argues that GE’s main aim is to explain why there is agglomeration of economic activity instead of dispersion, not why there is agglomeration in some places and not others. So CP may provide a successful EpHPE for the former explanandum, while failing for the latter. At any rate, the notion of EpHPE provides additional conceptual resources to appraise modelling controversies in economics.

4 Objectively Possible HPEs: Axelrod et al. on Social Networks and Cooperation

Evolutionary game theory models the evolution of cooperative behaviour as payoff-dependent replication of strategies in a population playing an indefinitely repeated prisoner’s dilemma (PD) game (for an overview, see Bowles and Gintis 2011). These modelling strategies involve substantial abstractions and idealizations. Yet they have centrally contributed to the understanding of cooperation, in particular regarding the possibility of non-kinship explanations of cooperation (e.g. Trivers 1971; Sugden 1986; Bowles and Gintis 2011). Perhaps most importantly, these models helped researchers to establish non-kinship-based mechanisms as genuine objective possibilities for the evolution of cooperation, by showing how from plausible initial assumptions, assuming standard economic ceteris paribus laws (mainly utility maximization), stable cooperation equilibria would emerge.

These models typically make substantial counterfactual assumptions, some of which directly contradicting what is known about the actual world. They are substantial in the sense that they cannot be justified as approximations or tractability assumptions, but instead represent explanantia or explananda known not to exist in the actual world. This indicates that these models do not establish epistemic possibilities, and thus do not provide EpHPEs. Instead, these models play important epistemic functions as ObHPEs.

Take for example Axelrod et al. (2003), who simulated the performance of populations on different network structures in an agent-based model in which groups of four play in a repeated PD. Groups either (i) are randomly paired in each round, (ii) are assigned a fixed location with fixed neighbours for the whole simulation, or (iii) are randomly paired at the beginning of the simulation but retain that pairing for the whole run. The authors find that populations do as well (in terms of average population payoffs) in (iii) as in (ii), and both do better than in (i). They conclude that geographic clustering is not necessary for social cooperation: “a persistent random network can support cooperation as well as
a geographic network” (Axelrod et al. 2002, 345).

Two aspects are noteworthy. First, when the authors wrote this article, actual humans were not participating in stable random networks. Arguably, social networking services — nascent around 2000 — have by today generated such kinds of networks, and perhaps the authors of this paper could foresee this, but it wasn’t reality then. Thus, at the time it couldn’t serve as the explanans of an actual phenomenon and even if stable random networks had never happened this would not have made the ObHPE less successful.

Second, the authors draw a modal conclusion: such networks can support cooperation as well as clustered ones. It is the possibility that matters — but of which kind?

Axelrod et al.’s model does not provide a candidate explanans that was epistemically possible given knowledge at that time. Rather, they aimed to show that a counterfactual explanans could produce cooperation, too. Previous research on cooperation had focused on clustered networks, and sometimes — at least implicitly — had assumed that clustered networks were necessary for establishing and maintaining cooperation. The authors’ goal, by modelling the objective possibility of distant network cooperation, was to contradict this implicit necessity claim.

While Axelrod et al.’s model therefore serves at least one epistemic function, not every model representing a counterfactual possibility does. Reasons for such a failure might be located in the justification of modal claims, an issue we decided to bracket in this paper (see footnote 1). Alternatively, such models fail pragmatically as an ObHPE when there is no demand for knowing such a possibility. For example, a necessity claim contradicted by such a possibility is not seriously considered or the possibility shown by the model falls outside of the modal scope of such contended necessity claims. This last case has received considerable attention in economic methodology.

For instance, one general and common criticism directed at general equilibrium modelling à la Arrow and Debreu (1954) is that it is overly ‘formalistic’ (Blaug 2002, 2003) or that it is just an empty piece of mathematics (Rosenberg 1992). One way to understand these worries is that the models provide true ObHPEs, but not of a relevant modality. That a general equilibrium is mathematically possible, critics hold, is irrelevant to the epistemic goals of economics. According to Blaug, the economic problem of a general equilibrium “has been transformed into a mathematical problem about a virtual economy, which is solved not by the standards of the economics profession, but by those of the mathematics profession (Walker 1997)” (2003, 148).

One crucial issue thus concerns the interpretation of the modal operator. If general equilibrium models were to provide economically possible ObHPEs, Blaug’s concern would at least be partially addressed. They might show, for instance, what sort of economic causes would need to be operational for a general equilibrium to occur. This might help to answer whether the general equilibrium is possible in a relevant way. But a mere mathematically possible ObHPE is not, according to Blaug, informative about economic processes. We can thus interpret such criticisms as concerning the relevance of specific
modal operators.\textsuperscript{6} Not everyone shares these particular concerns about general equilibrium models. For one, Hands (2016) acknowledges that even though a perfectly competitive economy does not actually exist (and could not exist according to our knowledge of the actual world), it still depicts a relevant possible world.

Notice that an actual perfectly competitive economy is an economy that has never existed nor ever will exist; it is a hypothetical economy with no monopolists, no oligopolies, no production by the government or non-profits, no brand names, free exit and entry into every industry, and a host of other features. And yet it is a possible world. It does not violate any laws of logic or nature, and a few sectors of real market economies even approximate it (like agriculture). (Hands 2016, 38, emphasis in original)

In other words, we could say that for Hands a perfectly competitive economy is epistemically impossible, but objectively possible. What general equilibrium models do is to explain how that objectively possible world could come about (see also Verreault-Julien 2017). Critically, according to Hands it was not an exercise in pure mathematics because many assumptions had an “economic interpretation”, i.e. they were consistent with stylized economic facts and background theory.

Both sides agree that the models establish true claims of objective possibility. But for some it is a ‘mere’ mathematical possibility (Blaug 2003), whereas for others it is something stronger like physical (or economic) possibility (Hands 2016). This suggests two potential points of contention concerning the epistemic value of general equilibrium modelling. One is about which sort of modal claims economics should seek to establish. Can claims of mathematical possibility ever be relevant for economics? Another is about the sort of claims general equilibrium models support. Do the models establish mathematical or economic possibility? As in the previous section, our goal here was not to take position in the debate over general equilibrium models, but simply to illustrate how the notion of ObHPEs illuminates the nature of the disagreement and can help normative appraisal.

5 Conclusion

Interpreting some modelling practices in economics as providing HPEs allows, in principle, to salvage their epistemic value from the requirement that good practices necessarily provide HAEs. However, this may seem to entitle economists to formulate any just-so story they wish. On what grounds could we possibly judge that a model didn’t provide an HPE? As we have argued, this first impression is misguided; there are criteria that enable to distinguish ‘good’ from ‘bad’ HPEs. We proposed a conceptual framework that made

\textsuperscript{6}The operator is also sometimes understood in terms of logical possibility: “With the existence results above, economists validated the logical possibility that prices alone could equilibrate all competing interests” (Athreya 2013, 68, emphasis in original). Again, critics may wonder why such logical possibility would be of epistemic value for economists.
two key distinctions: between HAEs and HPEs, and between EpHPEs and ObHPEs. That framework allows for the wider epistemic evaluation of economic modelling practices. Of course, this does not imply that most of these modelling practices actually have epistemic value — they might fail the criteria for good EpHPEs or ObHPEs. It also does not imply that it is an ‘efficient’ way of producing epistemic value (Northcott 2018). But in order to answer these questions, philosophers and methodologists of economics should pay more attention to the modal features of economics and their epistemic appraisal. Our modest goal was to show how this could possibly be done.

References


Sjölin Wirling, Ylwa, and Till Grüne-Yanoff. forthcoming. ‘Epistemic and Objective Possibility in Science.’


