

What economics can(not) tell?

Our reason is not like an indeterminably extended plane, the limits of which one can cognize only in general, but must rather be compared with a sphere, the radius of which can be found out from the curvature of an arc on its surface (...), from which its content and its boundary can also be ascertained with certainty. Outside this sphere (field of experience) nothing is an object for it (...).

Critique of Pure Reason
Immanuel Kant

This short essay aims at discussing certain limits of the economic theories' efficiency, particularly, of microeconomic modeling. Even though this field of science is sometimes compared in its method to physics (for probably the most well-known instance see [Friedman, 1953]), I believe that the obstacles it faces are to at least some extent idiosyncratic. Not solely restricted to the problem of data collection, the issue concerns economic models' compatibility with the material world, especially when they aim to explain "traditionally" non-economic phenomena, following the logic of economics imperialism. In most cases, one cannot easily say not only if taken assumptions are approximately satisfied, but also if the implications are not confounded by other active effects. What can models effectively say? And what they cannot? In what comes below, I will try to show that even if microeconomic¹ models cannot explain much as long as their own standards are treated seriously, they can nevertheless serve as benchmark frameworks of rationalizable human behavior. They do not reveal *actual* relationships between objects of study, but set up limits for their possibility, thus being specific instances of possible worlds. However, what really happens in each of these worlds is in a strict sense unknown. This unknowingness is an inevitable side-effect of modeling practice.

More than forty years after Deirdre McCloskey's *The Rhetoric of Economics*, the methodological situation within the discipline seems to not have changed too much, despite appearances. Obviously, new paradigms have emerged and some of them made it through to the core (especially behavioral economics), while

¹I restrict my interest to microeconomics due to its particular (when compared with, e.g., macroeconomics) tendency to apply methods of economics in domains of psychology, criminology, etc.

some previously disregarded methods (such as the usage of questionnaires) have become accepted. At the same time, the sole concern about verifiable behaviors without discussing agents' beliefs also seems to already belong to the past, mostly due to the inputs of game theory. However, the official modernist credo of economics McCloskey cites in her essay has stayed greatly intact. According to it, the main goal of science is prediction while only observable implications of a theory are of importance for its truth or falsity. Similarly, objectivity and quantitative measurement are praised, with subjectivity and qualitative assessments being exiled to other disciplines. Not to mention the surprising longevity of the falsificationist paradigm or positivistic approach according to which science shall not speak normatively about values or morality [McCloskey, 1983, p. 484-485]. Unsurprisingly, the number of equations per paper is known to have substantially grown over the last decades [Espinosa et al., 2012].

In fact, as McCloskey notes, economics has never fully followed those prescriptions, as no science could do it without landing in self-refutability. Among many theoretical conundra they lead to, she mentions the well-known objection against Popperian falsificationism (which auxiliary premises should be rejected if their conjunction with the main hypothesis is falsified?) or the doubtful consequences of the common claim that science's enterprise is to generate predictions [McCloskey, 1983, p. 487-488].

Speaking about falsification, in an ideal, positivist-shaped world, a theory is built by applying some assumptions, and its implications are verified empirically by available data. In a similar vein, Milton Friedman calls for testing a theory by its implications, not by assumptions [Friedman, 1953]. Let me leave aside whether his view should be considered more realistic [Mäki, 2009] or, on the contrary, more post-modern², than it is usually interpreted to be [McCloskey, 1983, p. 485]. Instead, one should rather ask about the status of economic theories whose implications are only possibly, yet not actually verifiable. This is not a trivial question, as economic models are usually constructed in a hugely hypothetical mode, relying on numerous "ifs". An important thing to ask about is then what reassures us that a model can be considered an appropriate representation of the world [Morgan, 2012, p. 32].

The difficulty grows more substantial when we realize that a big part of economic work is now done in theory, even if data possible to (dis)prove these a priori findings is not expected to be generated in any foreseeable future. This is especially the case for microeconomic theory in its extended application in the domains of political science, psychology or criminology³. However interesting the obtained results of such studies may be, the existence and character of equilibria, comparative statics, etc., a serious hardship they must face is that most of the costs and profits agents optimize over have a largely internal or unconscious character⁴. Collecting the data, not least for ethical reasons, is

²In its appeal to aesthetic criteria of a theory's choice.

³For some applications, see for example [Bénabou and Tirole, 2011], [Benoît and Dubra, 2004], [Ali and Lin, 2013].

⁴One among the differences with physics, at least its big parts, is that subjects of data-collecting are by themselves responsive to the fact of experimental inquiry. Sometimes simply

severely limited. Some help may be provided from the side of applied psychology methods; this, however, requires a metatheory, explaining how to measure those measurements themselves. As Nancy Cartwright argues, economic models do not resemble Galilean thought experiments due to their “overconstraint” nature. Contrary to modeling a falling body, economic models must include numerous assumptions (such as, for example, agents’ risk aversion, the timing of their decisions, (im)perfect information, etc.) already at the stage of building a theory, not only testing it [Cartwright, 2009, p. 48-50]. Without them, no results would be obtained. Along the same line goes Anna Alexandrova, dividing models’ assumptions into situation definers and derivation facilitators⁵. The latter, such as twice-differentiability of the utility function or statistical distributions in use, being of a secondary character for a situation modeled, simply make derivations easier (or possible at all). As theorists, we do not always have a good understanding of their possible empirical content or what their potential violation/relaxation would mean for the validity of the model’s findings [Alexandrova, 2006, p. 180-184].

Micro-theory is nevertheless thriving, and of course, it is not a problem in itself. Economists can build successful theories explaining relevant channels of the impact of various exogenous variables on human behavior, but the question remains what should be the measure of such success when neither implications nor assumptions are in all respects verifiable. Naturally, a theorist might well accept that logical consequences of given premises can be an interesting object of study, a discovery such as a newly observed asteroid [Li, 2023]. Such objects would be successful candidates for inhabitants of Gödel’s mathematical universe, in his view accessible not only by ways of abstract reasoning but also by means such as intuitive insight [Poreba, 2021]. This is a fair approach, but how to argue for its applicability to human behavior - the basic concern of economics?

What cannot also be easily said is which assumptions are sufficiently realistic, consequently constituting a good model. As Robert J. Shiller, quoting Jerome Bruner, notes (reminding thus in the economic context of the Duhem-Quine thesis), facts never speak for themselves [Shiller, 2017, p. 973]. The same is valid for the assumptions taken: they always require justification and apologetic argumentation. This, however, obviously requires either a higher instance or a kind of self-referential (transcendental?) grounding. In fact, even when a model’s predictions are testable, we cannot be in the deep sense sure it is not only because of some contingent confounders, not specified by the model, that this representation-to-world mapping holds. For similar reasons, Mary S. Morgan proposes to treat them rather as objects of investigation by experiment than as those of proofs and postulates [Morgan, 2012, p. 32-33]. If we, however, want to sustain models’ theoretical character, we need to investigate where they lead to: what they can say and what cannot.

Those thinking that economic theory is simply going round in circles and

running a regression across collected data might result in a selection bias and in consequence confound the results.

⁵In her opinion, the division distinguishes economics from Galilean thought experiments - in the latter, all assumptions are treated as situation definers [Alexandrova, 2006, p. 183-184].

repeating the old Cartesian problem of fundamental knowledge should not be misled. The quest is not for the undoubted truths providing cornerstones of all possible knowledge but for the applicability of formal theories to the explanation of human behavior, and their reasoned use. Shaun Gallagher and Enrico Petracca, following the works of Andy Clark, point out that neoclassical economics yields successful results where it considers situations with a relatively restricted (*scaffolded* in Clark's words) choice. If an agent's problem becomes vague enough, neat optimization formulas may not be sufficient to explain the whole variety of his or her behavior: including the factor of individual psychology and/or institutional analysis might be necessary [Petracca and Gallagher, 2020, p. 751-752]. Individual psychology (to dismiss a potential utilitarian reply) as irreducible to maximization of other sorts of utility and institutional analysis not giving up on institutions' normative content [Searle, 2015].

How then to avoid a gloomy conclusion that economists for sure might be reliable experts in IT, but no good theorists of human behavior [Markey-Towler, 2017, p. 8, fn. 7] or that an attempt to understand an agent's activity in mechanical terms must necessarily transform them into an almost-inhuman animal laborans? [Arendt, 2013, ch. 43-45]

An answer from within the microeconomic theory would be to apply its standards to itself and draw honest conclusions from it. Yes, we cannot "fully" know if a model's assumptions are justified or not. Yes, the less specific the context of a decision-making process, the greater the risk of ignoring important not-strictly-economic factors. Yes, we might not be able to test the identified channels or estimate the empirical parameters due to data limitations. What those weird mathematical objects, microeconomic models (of which one cannot be even sure to which domain they apply), can yield is, however, not strictly negative. By drawing conclusions from given premises, they provide possible worlds within which certain properties of described agents are held. Similarly, by ruling out some ex-ante feasible possibilities as impossible (contradictory or violating accepted axioms), they guarantee that in no possible world, certain scenarios occur.

Not too far from that view, Alexandrova [Alexandrova, 2008, 396-402] proposes to treat models as open formulae, frameworks for formulating economic hypotheses. What makes them different from existential claims is the application of free variables (unconstrained by quantifiers), enabling a smoother approach to difficulties of (not) satisfying a model's assumptions. In her view a model, in contrast to an existential claim, defines a situation *type* by enforcing a set of *possible* conditions, altogether identifying a causal link between A and B. Consequently, no material entity or relationship is directly postulated by such a theory. It is during the attempts to confirm a model empirically that a concrete situation fulfilling relevant criteria is pinned down. Thereby, free variables in use are filled and conditions set by the model might be restricted or expanded. In this way, a model sets a "recipe" for finding empirical content similar (yet not necessarily identical) to the relationships described by it. Confirmation is then an inevitable step of theory-building.

My focus here is however on models understood not as ready-to-fill formulae

allowing for drawing material hypotheses, but frameworks setting the limits of possibility through showing the (non)contradictory character of some claims. As said before, in many instances of microeconomic theory one could barely think of the findings' empirical verification. What they offer instead are limiting bounds for rationalizable human behavior. A neat example for the above claim may be provided by the inspection of a 1995 paper by Thomas Piketty, modeling a feedback loop between beliefs concerning social mobility and actual redistributive policies [Piketty, 1995]⁶. Even if agents' behavior can be shown to be explainable by a sort of utility maximization and beliefs updating under some conditions, and even if those processes converge to distributions somewhat similar to the ones observed empirically (as Piketty shows), this is no more than positing a hypothetical link (however interesting it can be, as in this case). The model's findings reveal that disparities between preferred tax rates and induced effort (low-high, high-low) *may* be mutually reinforcing and thus result in multiple equilibria. However, its successful construction does not determine the *actual* impact of given beliefs on preferred tax rates nor explains how "things really are". It reveals a set of non-contradictory scenarios by building a possible world in which certain characteristics are displayed. Whether that possible world overlaps with the actual one is, however, a different matter. It may well be that some crucial assumptions have been omitted or that the model's implications are confounded by some unobserved factors. Or even that those confounders are confounded as well and the model's results hold just by coincidence. What is truly interesting is proving the *possibility* of a given scenario (here, the reciprocal dynamics between the beliefs about social mobility and the preferred policies) or - quite differently - its impossibility as contradictory. Yet, we cannot know if the postulated link actually holds. From the perspective of the actual world, it remains in the sphere of uncertainty.

It should be added that, despite the terminological similarity, the proposition formulated here differs from Robert Sugden's idea of economic models as *credible worlds* [Sugden, 2000], [Sugden, 2009]. On the one hand, he also speaks of models as counterfactual worlds that reveal some possibilities [Sugden, 2009, p. 10, 16]. On the other, in his view, a model is mostly a means of abductive reasoning, from the similarity of effects (between the model and the real world) to the similarity of causes. The approach I propose is instead closer to the isolation of (causal) capacities as defended for example by Cartwright (models seen as tools used to build artificial environments in which causal links can be distilled and studied). In a much more akin tone speaks Till Grüne-Yanoff, who understands models as possible worlds that increase or decrease our confidence in some propositions concerning necessity or impossibility [Grüne-Yanoff, 2009]. For him, a model can be used to learn something about the actual world, if it presents a relevant possibility (1), contradicting an impossibility hypothesis held with sufficient confidence by some agents (2) [Grüne-Yanoff, 2009, p. 97]. My claim here is of a more ontological, less epistemological character. Namely,

⁶Although the topics discussed in Piketty's paper might appear more relevant for macroeconomists than for microtheorists, I refer to it due to a particularly microeconomic flavor in its modeling approach.

models as possible worlds are objects of interest because they express nontrivial modal properties, not necessarily due to their relation with agents' actual knowledge or beliefs. Naturally, these models can be later used for such belief-updating. However, by themselves, they also can express propositions about necessity or impossibility, not only change our views about them. Moreover, I believe that a model might still be informative even if it does not contradict any widely held impossibility claim, as demanded by Grüne-Yanoff. It might present (itself as) an alternative possible world, or in other words provide a competing, possibly simpler explanation of an investigated phenomenon.

At the same time, despite the appeal to the classical concept of possible worlds I allow myself to, some differences must be kept in mind. Its usage I propose above does not fully boil down to modal logic or semantics but also involves some axioms traditionally applied by microeconomic modeling, such as the ones accepted by the rational choice theory. The fact that empirical agents do not necessarily follow those axioms, as for example may be concluded from frequent violation of the Sure Thing Principle⁷, suggests it is a refinement of the concept of possible worlds. It departs from its traditional usage by imposing additional restrictions on it. One might notice that under this understanding the number of possible worlds grows vastly, maybe even to the power set of the "traditional" possible worlds' set. Which axioms are found relevant enough to hold and which not necessarily, is a more or less an arbitrary choice of a modeler; therefore, for most theories/models containing some axioms, say the Sure Thing Principle, we could construct a competing one without it. As a result, the domain of possibility understood in this way expands substantially⁸. Troubling as this idea may seem, at the end of the day it is the very construction of those possible worlds of consistently rational agents that has laid the groundwork for the emergence of behavioral economics. Without it, accounting for apparent breaches of so-conceived rationality, too frequent to treat them as mere anomalies, would not be viable.

Those possible economic worlds set up bounding frameworks, or in other words limits, for human action. Identified channels of influence that some elements of a model exert on the other or effects emerging out of strategic interactions between players hold as analytic relationships. They only constitute scaffolding inside of which agents can act. However, those frameworks themselves are mute, uninformative about their applicability. Therefore, anything not ruled out by them lies within the realm of possibility, and with it, uncertainty; each model thus unexpectedly opens a space of unknowingness. What is "really" taking place within it, economics cannot say. At least according to its own official standards of knowledge. Bringing the answer closer, and violating these rigid standards of modernist methodology, requires invoking what

⁷When agents play differently if not provided information about the true branch of a decision tree they are in: although in each feasible branch they would play the same, different, strategy [Shafir and Tversky, 1992].

⁸Possibly, its further inspection (maybe under a different name than that of *possibility*) would be a promising theoretical enterprise, at the moment however beyond the scope of this essay.

McCloskey has called the unofficial rhetoric of economics.

References

- [Alexandrova, 2006] Alexandrova, A. (2006). Connecting economic models to the real world: Game theory and the fcc spectrum auctions. *Philosophy of the Social Sciences*, 36(2):173–192.
- [Alexandrova, 2008] Alexandrova, A. (2008). Making models count. *Philosophy of Science*, 75(3):383–404.
- [Ali and Lin, 2013] Ali, S. N. and Lin, C. (2013). Why people vote: Ethical motives and social incentives. *American economic Journal: microeconomics*, 5(2):73–98.
- [Arendt, 2013] Arendt, H. (2013). *The human condition*. University of Chicago press.
- [Bénabou and Tirole, 2011] Bénabou, R. and Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, 126(2):805–855.
- [Benoît and Dubra, 2004] Benoît, J.-P. and Dubra, J. (2004). Why do good cops defend bad cops? *International Economic Review*, 45(3):787–809.
- [Cartwright, 2009] Cartwright, N. (2009). If no capacities then no credible worlds. but can models reveal capacities? *Erkenntnis*, 70(1):45–58.
- [Espinosa et al., 2012] Espinosa, M., Rondon, C., and Romero, M. (2012). The use of mathematics in economics and its effect on a scholar’s academic career.
- [Friedman, 1953] Friedman, M. (1953). The methodology of positive economics.
- [Grüne-Yanoff, 2009] Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis*, 70:81–99.
- [Li, 2023] Li, S. (2023). How to give an economic theory talk.
- [Mäki, 2009] Mäki, U. (2009). Unrealistic assumptions and unnecessary confusions: Rereading and rewriting f53 as a realist statement.
- [Markey-Towler, 2017] Markey-Towler, B. (2017). I, roboticus oeconomicus the philosophy of mind in economics, and why it matters. *Cambridge Journal of Economics*, 41(1):203–237.
- [McCloskey, 1983] McCloskey, D. N. (1983). The rhetoric of economics. *Journal of economic literature*, 21(2):481–517.
- [Morgan, 2012] Morgan, M. S. (2012). *Modelling as a Method of Enquiry*, page 1–43. Cambridge University Press.

- [Petraçca and Gallagher, 2020] Petraçca, E. and Gallagher, S. (2020). Economic cognitive institutions. *Journal of Institutional Economics*, 16(6):747–765.
- [Piketty, 1995] Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly journal of economics*, 110(3):551–584.
- [Poreba, 2021] Poreba, M. (2021). Gödel, Wittgenstein and the sensibility of Platonism. *Eidos. A Journal for Philosophy of Culture*, 5(1):108–125.
- [Searle, 2015] Searle, J. R. (2015). Status functions and institutional facts: reply to Hindriks and Guala. *Journal of Institutional Economics*, 11(3):507–514.
- [Shafir and Tversky, 1992] Shafir, E. and Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive psychology*, 24(4):449–474.
- [Shiller, 2017] Shiller, R. J. (2017). Narrative economics. *American economic review*, 107(4):967–1004.
- [Sugden, 2000] Sugden, R. (2000). Credible worlds: the status of theoretical models in economics. *Journal of economic methodology*, 7(1):1–31.
- [Sugden, 2009] Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis*, 70:3–27.