Poznań University of Economics and Business
Research Papers in Economics and Finance *Vol. 8, No. 2, 2024: 135–148*
https://ref.ue.poznan.pl *https://doi.org/10.18559/ref.2024.2.1902*

Underdetermination problem in methodology of economics

Damian Luty[1] **Rafał Iwański**[2]

Abstract

This paper explores the Duhem-Quine (DQ) problem and its impact on economic methodology, focusing on how the reliance on auxiliary assumptions complicates the testing and validation of theories. The DQ problem shows that no hypothesis is tested in isolation, as it depends on additional assumptions and background knowledge, making it challenging to pinpoint where errors lie. This issue is particularly relevant in economics, where complex models and assumptions about human behaviour play a significant role, and in finance, where the robustness of models is critical for decision-making under uncertainty. The paper highlights two key gaps: (a) the limited discussion of the DQ problem in economic methodology, and (b) the lack of alternative approaches to ensure rational methods in light of DQ. To address these issues, it proposes a multi-criterial framework for evaluating theories, emphasising consistency, diverse data, localised testing, comparing models and varying assumptions systematically. Using examples such as housing market models and the Ultimatum Game, the paper illustrates how addressing the DQ problem involves avoiding arbitrary changes to assumptions while adopting clear, rational strategies. By providing a stronger methodological foundation, this approach enhances the reliability of economic and financial theories, improving their influence on policy-making and practical applications.

Keywords

- Duhem-Quine problem
- economic methodology
- methodological rationality

Article received 4 November 2024, accepted 11 December 2024.

Suggested citation: Luty, D. & Iwański, R. (2024). Underdetermination problem in methodology of economics. *Research Papers in Economics and Finance*, *8*(2), 135–148. https://doi.org/10.18559/ref.2024.2.1902

[1] Poznań University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland, corresponding author: damian.luty@ue.poznan.pl

[2] University of Szczecin, al. Piastów 40b, 71-064 Szczecin, Poland, rafal.iwanski@usz.edu.pl

Introduction

The notion of rationality in the context of economic theory is usually presented as an exceptional and fundamental principle, especially in mainstream (monetarist/neoclassical) economic theory. The principle of economic rationality, however, is far from a homogenous concept, especially since the emergence of behavioural economics (Kahneman et al., 1986) and the concept of bounded rationality (Simon, 2008). To put it generally and somewhat synthetically, to speak of economic rationality means speaking of *choice* with respect to means allowing for maximalisation of utility for the economic agent; a *rational choice* is that which yields the highest utility for the agent with respect to the agents preferences about scarce resources (see: Jones, 2021; Li, 2020; Sen, 1977; Smith, 2007). From this perspective "Rational people systematically and purposefully do the best they can to achieve their objectives, given the available opportunities" (Mankiw, 2008, p. 6). Economic rationality conceived as the art of making choices[3] was, as is well known, formally modelled by Arrow (1963, 1974).

Less known thematisation of the notion of rationality regarding economics has to do with methodological and philosophical considerations about ways in which economic theories fulfil criteria of scientific rationality. This perspective conceives rationality as a property of economics considered as a scientific enterprise, while the former perspective views rationality as a property of economic agents. This type of rationality will be referred to as "methodological rationality" in this article. Firm belief in the rationality of the whole enterprise seems to be a necessary prerequisite in making epistemic and, further, policy-making analysis, modelling or predictions (Newton-Smith, 1981). As such, rational behaviour at the methodological level would require providing reasons as to why such-and-such, e.g. predictions are not *ad hoc*. To claim otherwise would mean to sanction the idea that, fundamentally, everything goes regarding economic theorising (Kuorikoski & Marchionni, 2024). Even though it is widely recognised that "pure" rationality at the methodological level, understood as the capability to reconstruct in terms of logic ways in which hypotheses are formulated and connected with empirical evidence, is impossible (Feyerabend, 1993; Kuhn, 1962), to surrender to arbitrariness as the methodological guiding principle seems to be a cost that is too high – especially with the general rising tide of pseudoscience.[4]

[3] It is worth noting that this concept is also applied in other social sciences (e.g. political science, international relations). However, some authors emphasise its significant limitations, both in economics and in other social sciences (Mearsheimer & Rosato, 2023).

[4] By "pseudoscience" we mean some set of beliefs falsely being presented or considered as based on the scientific method.

This paper addresses one of the main challenges to the methodological rationality of economics (as well as other scientific disciplines): the so-called underdetermination of theory by data, or "the underdetermination problem", or "the Duhem-Quine Problem" (DQ). As such, the work presented here is of conceptual nature. The research gap identified in the presented context consists of two parts: (a) the problem is not well known in the literature on economic methodology, so it is important to highlight this issue; (b) no alternative concept of methodological rationality, aside usual organisational standards within scientific institutions, has been proposed in the light of the DQ problem. The first gap is filled in this paper by presenting the DQ in economics; the second gap is filled in section 4 by proposing a multi-criterial concept of (general and synthetic) methodological rationality for economic theorising.

1. Research methodology

In this paper the critical review method is utilised (Grant & Booth, 2009; de Klerk & Pretorius, 2019). The critical review method involves systematically analysing, synthesising and evaluating existing literature to provide a comprehensive understanding of a specific research topic. This methodology is particularly valuable for identifying gaps, trends and debates within a field, offering a solid foundation for further research. By critically engaging with sources, the method goes beyond summarisation to assess the validity, reliability and relevance of existing work. Critical engagement with sources amounts to a synthetic presentation of the DQ problem in existing literature on methodology of economics and formulating an addition to it by formulating potentially fruitful methodological hypotheses. Data set consist in methodological and philosophical articles dedicated to the DQ problem – both at the general level and in economics. The keywords "Duhem AND Quine", "Duhem AND Quine AND problem", "Duhem", "Quine", "underdetermination" were used for researching the literature in Scopus and Google Scholar databases. The literature was chosen based on the clarity and thoroughness of characterisations of concepts and topics relevant for the paper, i.e. general DQ problem, DQ problem in economics, as well as examples of DQ problem in economics. Literature from no specific time period was considered, mainly because the DQ problem is neglected in methodological literature in economics. The Scopus database yields 7 items, when the research is limited to economics and finance literature. This is the reason why additional sources were used: books and articles evidenced in the philpapers.org, a website dedicated to aggregating information on philosophical papers. Ultimately, 38 items were selected in the process and subsequently subjected to critical analysis and synthesis.

2. Literature review

Let us limit this section to literature relevant to the DQ problem in economics. The seminal paper regarding this topic was presented by Cross (1982). Cross explores the implications of the Duhem-Quine (DQ) thesis for economic methodology by examining its influence on the Keynesian revolution. He argues that the DQ thesis – asserting that no hypothesis can be tested in isolation – poses challenges for economists seeking empirical validation for macroeconomic constructs. This early exploration underscores the need for pragmatic strategies in economic methodology to navigate the interconnectedness of economic models. V. Smith (1994) focuses on the rise of experimental economics and how laboratory experiments have introduced new ways to test economic theories. He discusses the inherent reliance on auxiliary hypotheses, such as the *ceteris paribus* conditions and participant behaviour assumptions, which align with the DQ thesis. Smith emphasises that while laboratory settings can control variables, they do not eliminate the complexity of auxiliary assumptions. He advocates for transparency in experimental designs in order to improve the robustness of economic hypotheses. Sawyer and Sankey (1997) delve into the philosophical underpinnings of the DQ thesis, focusing on its relationship with scientific realism. They propose that scientific theories should be evaluated based on their explanatory power and ability to generate empirical predictions. Their work contributes a broader philosophical perspective on how the DQ thesis applies across disciplines, including economics.

Starmer (1999) critically examines the reliability of experimental economics, questioning whether the methods employed by experimental economists are sufficiently rigorous to justify their findings. Echoing the concerns raised by the DQ thesis, Starmer highlights the role of auxiliary assumptions in interpreting experimental outcomes. He raises ethical and methodological questions about the validity of controlled experiments in capturing real-world economic phenomena, advocating for more scrutiny of experimental methods. Boylan and O'Gorman (2003) revisit the DQ thesis in the context of economic methodology, emphasising a pragmatic approach to navigating its challenges. Jones (2012) analysed the DQ thesis with a focus on economic methodology. He critiques two foundational assumptions in Quine's depiction of science: the interconnectedness of scientific statements and the flexibility of auxiliary hypotheses. By challenging these assumptions, Jones argues for localised testing and constraints on modifying auxiliary hypotheses to preserve theoretical integrity. His analysis contributes to a nuanced understanding of the DQ problem in economics, offering practical approaches to address it. Mäki (2013) examines the role of contested models in economics, framing them as a reflection of the discipline's methodological challenges. He ties these challenges to the DQ thesis, arguing that economic models often compete for explanatory dominance,

relying on different sets of assumptions and auxiliary hypotheses. Mäki suggests that rather than resolving these contests outright, economists should embrace their diversity as a source of progress. He emphasises the importance of evaluating models based on coherence, empirical adequacy and practical applicability.

In every case, the emphasis was understandably placed on ways to mitigate the DQ problem. However, the main issue raised by DQ – the challenge to scientific rationality – remains unaddressed. In the literature reviewed it is clearly argued that from the methodological standpoint the DQ requires a response at the level of reflections on economic theory; however, the notion of rationality, viewed as a property of economic scientific discourse, is never tackled. The theoretical stake is high, given the importance of the notion of scientific rationality. As mentioned above, this prompts two research gaps filled by this paper: (a) the problem is not well known in literature on economic methodology, so it is important to highlight this issue; (b) no alternative concept of methodological rationality, aside usual organisational standards within scientific institutions, has been proposed in the light of the DQ problem.

3. General Duhem-Quine problem

The underdetermination thesis refers, first and foremost, to a relation between scientific theory (or the whole body of scientific knowledge) and empirical evidence (see: Psillos, 1999, p. 156). Originally, the thesis was formulated by Duhem (1954) in his analyses of physical science:

> The prediction of the phenomenon, whose nonproduction is to cut off debate, does not derive from the proposition challenged if taken by itself, but from the proposition at issue joined to that whole group of theories; if the predicted phenomenon is not produced, the only thing the experiment teaches us is that among the propositions used to predict the phenomenon and to establish whether it would be produced, there is at least one error; but where this error lies is just what it does not tell us. (p. 185)

That is, in scientific endeavours when a hypothesis is purportedly tested, it is never tested in isolation; it is always supported by different hypotheses or theories. If a hypothesis is rebutted we may claim that the fault lies in the supporting hypotheses – usually called the auxiliary hypotheses and background knowledge. In Duhem's philosophy of science, this underdetermination is an empirical matter to be decided by evaluating theories case by case. For over half a century, the problem was generalised by Quine (1951) to other types of knowledge:

> The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure math-

> ematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. But the total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in the light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole. (pp. 42–43).

Since then the thesis is known as the "Duhem-Quine thesis" (DQ) (Grünbaum, 1960). Roughly, DQ's main insight may be captured in the following way: "When we assert that scientific theory choice is underdetermined by evidence, we mean that evidence by itself cannot direct a scientist to accept or reject a theory" (Turnbull, 2018, p. 2). The acknowledgment of the DQ thesis had some large consequences in philosophy and methodology of science.

DQ undermines strict and pure accounts of scientific rationality. Most famously, DQ undercuts the logic of scientific justifications presented in Popper's falsificationism (Popper, 1959). According to Popper, a theory is scientific only if it is, in principle, possible to establish its falsity, i.e. it is not immune to possibly discrediting data. This forms the backbone of Popper's claims about the logic of scientific rationality – the correct logic in science is the logic of *modus tollendo tollens*. A rational scientist should always accept an argument which has the following form:[5]

$$H_0 \rightarrow O \wedge \neg O \rightarrow \neg H_0$$

where H_0 is the hypothesis tested and O is the observation. However, owing to DQ, this formula is inadequate: H_0 is never tested in isolation (what is actually tested is the conjunction of H_0 and auxiliary hypotheses and elements of background knowledge) and if recalcitrant experience is unfavourable for H_0, other parts of the theory or system can be blamed. The point is that, according to Quine, this can be done arbitrarily; if H_0 is supposedly refuted by evidence, we may try to save it by modifying the auxiliary hypotheses.

4. Duhem-Quine problem in economics

In methodology of economics, DQ was recognised in the following ways. Some authors emphasize DQ's meaning, some downplay it, others remain agnostic to its relevance. Mäki (2013) generally agrees that the DQ is pressing in economics and

[5] The formula in propositional logic is: $[(p \rightarrow q)\neg q] \rightarrow \neg p$.

that it amounts to a situation in which a decision between economic models cannot always be settled by empirical means. Hausman (2007) elaborates on this point:

> First, the complexity of human behaviour requires the use of numerous initial conditions and strong simplifying assumptions. Some of these restrictions may actually be false (such as the infinite divisibility of commodities), some of these assumptions may be logically unfalsifiable (such as the assumptions of eventually diminishing returns), while still others may be logically falsifiable but practically unfalsifiable (such as the completeness assumption in consumer choice theory). (p. 191)

An appraisal of DQ-problem in experimental economics can be found in (Guala, 2005), following V. Smith (1994):

> All tests of a theory require various auxiliary hypotheses that are necessary in order to interpret the observations as a test of the theory. These auxiliary hypotheses go under various names: initial conditions, ceteris paribus clauses, background information, and so on. Consequently, all tests of a theory are actually joint tests – that is, a test of the theory conditional on the auxiliary hypotheses. (p. 127)

To give two examples of DQ in economics, let us refer to McMaster and Watkins (2006) as well as Jones (2012).

Example 1 (McMaster & Watkins, 2006)
A substantial body of research employs econometric methods to analyse urban housing market data. Typically, these models define house prices based on physical property features, neighbourhood quality and proximity to the central business district (CBD). The framework assumes that metropolitan areas function as unified urban housing markets, where households select their homes by balancing location accessibility and housing quality within their budget constraints. In this setup, the coefficient associated with the distance from the CBD reflects the slope of the bid-rent gradient, all else being equal. Even if the distance variable shows an insignificant or positive relationship, the access-space trade-off remains empirically supported within this analytical framework.

Example 2 (Jones, 2012)
Another case is the testing of the ultimatum game, where the primary hypothesis under examination is that of self-interest. While the ultimatum game appears to provide a clear and straightforward test of this hypothesis, its validity depends on a broad set of auxiliary assumptions, complicating its interpretation as a definitive measure of self-interest. These auxiliary assumptions include factors such as whether the payoffs are sufficient to motivate participants, whether subgame perfect equilibrium is the appropriate equilibrium concept, whether irrelevant changes in payoffs influence behaviour, and whether variables like players' characteristics (e.g., gender, nationality, culture), or their experience, impact performance.

Over time, these auxiliary hypotheses have been questioned, either to reinterpret the findings of the ultimatum game or to evaluate the robustness of its results.

Just in these examples the threat of arbitrality in economic theory is visible: presuppositions present in the abovementioned cases – considered as crucial devices of explaining or making predictions or adequately representing the situation – are proven as having loose connection to the phenomena considered. Other examples of the underdetermination problem can be found in Cross (1982), Sawyer and Sankey (1997), Starmer (1999), Hands (2001), McGovern (2006) and Bardsley et al. (2010).

One way to deal with the DQ problem is to question the Quinean view on science, as proposed by Jones (2012). His critique of the DQ thesis can be outlined as follows. He argues that two key assumptions in Quine's depiction of science need to be challenged. The first is the concept of science as an interconnected web of statements, and the second is the flexibility in altering auxiliary hypotheses. These assumptions are central to the DQ problem. If science is not as interconnected as Quine suggests, localised testing becomes a plausible approach. Similarly, if auxiliary hypotheses are not easily modifiable, it becomes harder to adjust them solely to preserve a theory. While it is true that experiments depend on auxiliary hypotheses to test the main hypothesis, it is equally valid that a significant portion of experimental science involves verifying these auxiliary hypotheses.

Another way of mitigating the DQ problem is to provide a set of a general list of possible courses of action, which is to be consulted whenever doubts about auxiliary hypotheses arise. While the notion of purely logical methodological rationality usually took a mono-criterial form (referring to a specific logical law), post-DQ methodological rationality in the context of economy may take a multi-criterial form and be understood as a meta-frame of reference. For example:

1. **Model Consistency**: When a model reliably predicts real-world outcomes across different scenarios, it supports the credibility of its underlying hypotheses, even if isolating auxiliary assumptions proves challenging.
2. **Data Pluralism**: Evaluating economic theories using diverse datasets or contexts. Consistent results across varying conditions strengthen the argument that findings are not solely influenced by flawed auxiliary assumptions.
3. **Localisation**: Conducting focused testing of economic theories within smaller subsystems. This approach assumes that certain auxiliary assumptions remain constant, enabling more precise hypothesis evaluation.
4. **Comparison**: Instead of testing a single hypothesis, economists can evaluate multiple models or hypotheses to identify which one best aligns with the data. Bayesian methods are particularly useful for updating hypothesis probabilities based on new evidence while incorporating prior beliefs and auxiliary assumptions.

5. **Variations**: By systematically adjusting auxiliary assumptions, economists can gauge the extent to which the main hypothesis depends on them. If a hypothesis remains valid across a range of plausible conditions, confidence in its robustness increases.

This list is clearly tailored for a specific discipline; one could view this as something unfortunate, since no absolute reference frame of rationality is provided. However, this is a lesson that needs to be learned from DQ: an absolute/pure notion of rationality is impossible. At the same time, this is no final reason to abandon the notion of methodological rationality.

The third way of dealing with the DQ problem is to utilise the fact that we can distinguish different versions of DQ. The first version can be called "Duhemian underdetermination": contrastive underdetermination, i.e. underdetermination which arises due to empirical equivalence of competing theories or models (Stanford, 2023). The second version is the weak DQ: "single descriptive propositions are never tested in isolation; rather, empirical testing presupposes complexes or systems of sentences" (Boylan & O'Gorman, 2003, p. 12). The third version is the moderate DQ which adds the possibility of modifying *some* parts of the mentioned conjunction in order to save the tested proposition from recalcitrant evidence: "No *descriptive* statement can be individually falsified by evidence, whatever the evidence may be, since adjustments in the rest of the system can always be devised to prevent its *falsification*" (Hesse, 1970, p. 195). The fourth version is the strong DQ, which is expressed in the claim that "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system" (Quine, 1976, p. 60). This means that, presumably, we can defend a hypothesis if we make changes in any part of the entire system of knowledge.

Now, the Duhemian underdetermination does not pose a threat to economic theories since it is usually not the case that the choice between models or theories is about perfectly equal competitors: there are reasons to pick one theory over the other, even if it is because of values like simplicity or mathematical elegance, which are important in practical applications. The weak DQ, arguably most important in the context of the examples mentioned in this paper, was addressed in Boylan and O'Gorman (2003). They claim that

> practising scientists tend to divide their theories into high-level and low-level parts and they frequently hold that the lower-level is better corroborated or confirmed than the higher-level. Consequently, if a scientific theory is falsified by empirical evidence, the scientists, quite correctly, tend to locate the responsibility in the less confirmed parts of the theory. Without prejudice to the complexities of confirmation theory, weak Duhem-Quine theorists have no difficulty either with this pragmatic strategy. (p. 14)

According to these authors, the weak DQ is harmless, while the moderate version of DQ poses potentially a large issue for economics. However, we argue with

this claim. Auxiliary hypotheses cannot be arbitrarily altered to uphold the main hypothesis, as they often represent well-established theories with strong predictive accuracy. Changing these hypotheses without justification risks replacing a thoroughly validated hypothesis with one that has little or no empirical support. While a better hypothesis might theoretically exist, there is no guarantee of finding one, and it is often unlikely. This idea is clearly illustrated in economic experiments, where some auxiliary hypotheses are not even considered for revision. For example, the shape of a researcher's glasses is not seen as having any meaningful impact on a proposer's behaviour in the Ultimatum Game (see: Jones, 2012). Additionally, one could retort to some multi-criterial account of methodological rationality and claim that the possibility of changing some part of the system is actually good news, since to change that part may mean to improve the system; but only if this decision is guided, in a non-dogmatic way, by rationally grounded considerations.

Conclusions

The Duhem-Quine (DQ) problem presents a significant challenge to the methodological rationality of economics and other scientific disciplines. It underscores the inherent difficulty in isolating and testing individual hypotheses, as they are always embedded within a network of auxiliary assumptions and background knowledge. In economics, this interdependence complicates the validation of theories and models, as seen in the examples provided. Addressing the DQ problem requires rejecting arbitrary adjustments to auxiliary hypotheses, given their foundational role in supporting robust predictions and their alignment with broader theoretical frameworks. Furthermore, nuanced approaches, such as multi-criterial methodological rationality, offer a path forward. By evaluating models through criteria such as consistency, data pluralism, localisation and systematic variation, economists can improve the robustness of their hypotheses while maintaining methodological rigor. The distinctions between different versions of the DQ problem – ranging from weak to strong forms – highlight varying degrees of challenges to scientific inquiry. While weaker forms of the DQ problem may be pragmatically managed by prioritising corroborated components of theories, moderate and stronger forms demand careful, rationally grounded approaches to theory adjustment. Ultimately, the ability to refine systems of knowledge rationally, rather than arbitrarily, allows for progress in economic methodology. The solutions proposed in this paper aim to balance theoretical flexibility with the need for methodological consistency, ensuring that scientific advancements remain grounded in rational principles.

Building on the insights presented in this study, several avenues for future research can be pursued to deepen the understanding of the DQ problem in economics and enhance methodological approaches to address it effectively:

1. **Empirical Exploration of Auxiliary Hypotheses in Economics**

   Future research could empirically investigate how auxiliary hypotheses function in specific economic models. For example:
   a) What factors influence the choice and modification of auxiliary hypotheses in economic theorising?
   b) How often are auxiliary hypotheses revised, and under what conditions are such revisions considered justified?
   c) Case studies could focus on high-stakes areas like macroeconomic policy models or experimental economic games.

2. **Development of Practical Multi-Criterial Frameworks**

   While this study proposes a conceptual multi-criterial approach to methodological rationality, further work is needed to operationalise this framework. Researchers could:
   a) Develop tools or guidelines to apply criteria like model consistency, data pluralism and localisation systematically in economic research.
   b) Evaluate the practical impact of such frameworks on decision-making in theory selection or model validation.

3. **Integration with Machine Learning and Computational Methods**

   The growing use of machine learning and computational models in economics offers an opportunity to revisit the DQ problem:
   a) How do machine learning models, which often rely on complex auxiliary assumptions, navigate issues of underdetermination?
   b) Can computational tools be used to systematically evaluate the robustness of auxiliary hypotheses and their role in theory testing?

4. **Policy Implications of the DQ Problem**

   Future work could explore the practical implications of the DQ problem for policy-making, particularly in economics:
   a) How does underdetermination influence the reliability of economic predictions used to guide policy decisions?
   b) Can multi-criterial methodological frameworks improve the robustness and transparency of economic policy recommendations?

5. **Educational Strategies for Methodological Rationality**

   Finally, research could focus on incorporating an understanding of the DQ problem and methodological rationality into the education of economists:

a) What pedagogical approaches can help future economists better recognise and address underdetermination in their work?
b) Can case studies of the DQ problem in practice enhance the teaching of economic methodology?

By pursuing these lines of inquiry, future research can continue to refine the theoretical and practical approaches to the DQ problem, enhancing the rigor and reliability of economic science.

References

Arrow, K. J. (1963). *Social choice and individual values*. Yale University Press.

Arrow, K. J. (1974). *Essays in the theory of risk-bearing*. North-Holland Publishing Company.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2010). *Experimental economics: Rethinking the rules*. Princeton University Press. https://doi.org/10.1515/9781400831432

Boylan, T. A., & O'Gorman, P. F. (2003). Pragmatism in economic methodology: The Duhem-Quine thesis revisited. *Foundations of Science*, *8*(1), 3–21. https://doi.org/10.1023/A:1022417025502

Cross, R. (1982). The Duhem-Quine thesis, economic methodology, and the Keynesian revolution. *Journal of Economic Issues*, *16*(1), 119–133.

Duhem, P. (1954). *The aim and structure of physical theory*. Princeton University Press. https://doi.org/10.1515/9780691233857

Feyerabend, P. (1993). *Against method* (3rd ed.). Verso.

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Grünbaum, A. (1960). The Duhemian argument. *Philosophy of Science*, *27*(1), 75–87. https://doi.org/10.1086/287714

Guala, F. (2005). *The methodology of experimental economics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511614651

Hands, D. W. (2001). *Reflection without rules: Economic methodology and contemporary science theory*. Cambridge University Press. https://doi.org/10.1017/CBO9780511612602

Hausman, D. M. (2007). *The philosophy of economics: An anthology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511819025

Hesse, M. (1970). Duhem, Quine and a new empiricism. In G. Vesey (Ed.), *Knowledge and necessity* (pp. 191–209). Macmillan. https://doi.org/10.1007/978-1-349-86205-4_11

Jones, M. K. (2012). What to do with a problem like Duhem-Quine? *Revue de Philosophie Économique*, *13*(1), 79–102. https://doi.org/10.3917/rpec.131.0079

Jones, M. K. (2021). The concept of rationality in introductory economics textbooks. *Citizenship, Social and Economics Education*, *20*(1), 37–47. https://doi.org/10.1177/2047173421994333

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, *76*(4), 728–741.

de Klerk, W., & Pretorius, J. (2019). Guideline for conducting critical reviews in psychology research. *Journal of Psychology in Africa*, *29*(6), 645–649. https://doi.org/10.1080/14330237.2019.1691793

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Kuorikoski, J., & Marchionni, C. (2024). Economic models and their flexible interpretations: A philosophy of science perspective. *Journal of Economic Methodology*, 1–8. https://doi.org/10.1080/1350178X.2024.2336048

Li, C. (2020). The rationality principle as a universal grammar of economic explanations. *The Journal of Philosophical Economics: Reflections on Economic and Social Issues*, *13*(2), 58–80. https://doi.org/10.46298/jpe.10735

Mäki, U. (2013). Contested modeling: The case of economics. In U. Gähde, S. Hartmann, & J. H. Wolf (Eds.), *Models, simulations, and the reduction of complexity* (pp. 87–106). de Gruyter. https://doi.org/10.1515/9783110313680.87

Mankiw, N. G. (2008). *Principles of economics* (5th ed.). South-Western College Publishing.

McGovern, P. (2006). Methodological issues in economics. *Journal of Economic Methodology*, *13*(1), 1–19.

McMaster, R., & Watkins, C. (2006). *Economics and the housing market: An econometric perspective*. Routledge.

Mearsheimer, J. J., & Rosato, S. (2023). *How states think: The rationality of foreign policy*. Yale University Press. https://doi.org/10.12987/9780300274967

Newton-Smith, W. H. (1981). *The rationality of science*. Routledge. https://doi.org/10.4324/9780203046159

Popper, K. (1959). *The logic of scientific discovery*. Hutchinson.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. Routledge.

Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, *60*(1), 20–43. https://doi.org/10.2307/2181906

Quine, W. V. O. (1976). *The ways of paradox and other essays*. Harvard University Press.

Sawyer, R. K., & Sankey, H. (1997). Underdetermination in economics: The Duhem-Quine thesis. *Economics and Philosophy*, *13*(1), 1–23. https://doi.org/10.1017/S0266267100004272

Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs*, *6*(4), 317–344.

Simon, H. A. (2008). *Models of bounded rationality: Empirically grounded economic reason*. MIT Press.

Smith, A. (2007). *An inquiry into the nature and causes of the wealth of nations*. Liberty Fund.

Smith, V. L. (1994). Economics in the laboratory. *Journal of Economic Perspectives*, *8*(1), 113–131. https://doi.org/10.1257/jep.8.1.113

Stanford, K. (2023). Underdetermination of scientific theory. In *Stanford encyclopedia of philosophy*. Stanford University.

Starmer, C. (1999). Experiments in economics: Should we trust the dismal scientists in white coats? *Journal of Economic Methodology*, *6*(1), 1–30. https://doi.org/10.1080/13501789900000001

Turnbull, W. (2018). Underdetermination and the philosophy of economics. *Philosophy of Science Archive*, *2*(3), 1–20.