PhD Programme in Economics

Universities of Florence, Pisa, and Siena

XXXVI cycle

PhD Coordinator: Prof. Simone D'Alessandro

Essays on Economic Modeling and Expert Testimony in Antitrust Litigation

Candidate

Edoardo Peruzzi

Supervisor

Prof. Nicola Giocoli

Academic Year

2023/2024

ESSAYS ON ECONOMIC MODELING AND EXPERT TESTIMONY IN ANTITRUST LITIGATION

EDOARDO PERUZZI

Thesis submitted for the degree of Doctor of Philosophy in Economics
PhD Programme in Economics of the Tuscan Universities
University of Florence | University of Pisa | University of Siena

May 2024

Edoardo Peruzzi: *Essays on Economic Modeling and Expert Testimony in Antitrust Litigation*,, Thesis submitted for the degree of Doctor of Philosophy in Economics , © May 2024

SUPERVISOR:
Nicola Giocoli

Walk in silence
Don't walk away, in silence.

— Atmosphere - Joy Division

Dedicated to my family.

PUBLICATIONS

The author acknowledges permission to use previously published material in this dissertation. The provenance of the chapters is as follows:

Chapter 1. In part, from Paseyro Mayol, Joaquin and Edoardo Peruzzi. 2023. "The Profit Paradox: How Thriving Firms Threaten the Future of Work, Jan Eeckhout. Princeton, NJ: Princeton University Press, 2021, viii 327 pages". *Economics & Philosophy*, 39 (2), 338-343. DOI: 10.1017/S0266267123000044.

In part, new.

Chapter 2. Peruzzi, Edoardo, and Gustavo Cevolani. 2022. "Defending De-idealization in Economic Modeling: A Case Study". *Philosophy of the Social Sciences* 52 (1–2): 25–52. DOI: 10.1177/00483931211049759.

Chapter 3. Peruzzi, Edoardo. 2023. "Models on Trial: Antitrust Experts Face Daubert Challenges". *Journal of Economic Methodology*. 30 (4), 337-351. DOI: 10.1080/1350178X.2023.2267052.

Chapter 4. Unpublished material.

Chapter 5. Unpublished material

Appendix. Unpublished material.

ACKNOWLEDGMENTS

I wrote the chapters in this PhD dissertation in three different countries – Italy, Finland, and the United States – and have benefited from the help of many people. I have discussed various parts of the dissertation with Emrah Aydinonat, Gustavo Cevolani, Gianluca Damiani, Fabrizio Esposito, Kobi Finestone, Francesco Guala, Conrad Heilmann, Andrea Incerpi, Kevin Hoover, Kenan Huremović, Lisa Kinspergher, Teemu Lari, Uskali Mäki, Thiago D. Oliveira, Joaquín Paseyro, Nadia Ruiz, Giovanni Tuzet, and Jesús Zamora Bonilla.

Thierry Kirat and Steven Medema have meticulously reviewed the whole thesis and offered invaluable insights and suggestions.

Two visiting periods at the TINT Centre for Philosophy of Social Science, University of Helsinki, and Center for the History of Political Economy, Duke University, have been crucial to develop the arguments of the thesis. I am grateful to both institutions for providing me with the opportunity to present my work.

My supervisor, Nicola Giocoli, has been a constant, attentive and inspiring presence during the PhD. His contribution to this thesis has been fundamental, and I am profoundly grateful for everything he has done for me over the past four years.

The Appendix relied on data from the website *Daubert Tracker* (www.dauberttracker.com). I sincerely thank Myles Levin for granting access to the data.

I would like to extend my heartfelt gratitude to my family, friends from the Satta Messa, colleagues from the XXXVI PhD cycle, the House of Freedom in Durham, NC, and all the remarkable individuals worldwide whom I have had the privilege to meet during my doctoral journey.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

EC	European Commission
FTC	Federal Trade Commission
DOJ	Department of Justice
HMG	Horizontal Merger Guidelines
IO	Industrial Organization
TFEU	Treaty on the Functioning of the European Union

1

INTRODUCTION

To be frank, the law and economics of market power is a confusing mess.

— William E. Smith, District Judge[1]

Is the expert honest? At very best, probably as honest as is possible in a process in which truth is sought by the vigorous presentations of opposing views, and where any admission by one side is heavily overemphasized by the other side.

— George J. Stigler[2]

Scientific reasoning plays a crucial role in legal proceedings. DNA testing often serves as a key piece of evidence in criminal cases, while pharmacological studies come into play in toxic tort litigation, where the central question revolves around whether exposure to certain chemicals causes specific diseases. Additionally, psychologists and psychiatrists routinely testify about whether a murderer suffers from a mental disorder.

Economics is no exception. In fact, economists regularly provide expert testimony in various legal areas, ranging from antitrust enforcement to employment discrimination cases. Both economic theory and econometrics find practical application in the courtroom: calculating damages resulting from copyright violations, determining whether wages were lower due to gender discrimination, and predicting the impact of a merger on market prices – these are all issues for which economic expertise is essential.

This work examines the interaction between economic experts and the courts of justice in the United States, with a specific focus on antitrust law and its enforcement through lawsuits filed by public authorities, including the Federal Trade Commission (FTC) and the US Department of Justice (DOJ) Antitrust Division, and private parties who have been harmed by antitrust violations.

Economists have been actively engaged in advisory roles for policy-making purposes, employing macroeconomic models to guide our understanding of inflation, GDP growth, unemployment, and more. This interaction between economic experts and policymakers has generated extensive literature.[3] Auction models serve as another example

1 *In re Loestrin 24 Fe Antitrust Litig. 433 F. Supp. 3d 274 (D.R.I. 2019)*

2 *Memoirs of an unregulated economist*, Stigler (1988, p. 133)

3 See, for instance, the 2023 special issue of the *European Economic Review* (Vol. 159, October 2023) on the history of macroeconomics edited by Roger Backhouse and Beatrice Cherrier and the references therein.

of economic theory transitioning from academia to policy-making. Since the late 1990s, both US and European governments have employed game-theoretic auction models to distribute electromagnetic spectrum licenses to private firms. This approach allowed governments to allocate spectrum licenses in a way that fulfilled specific objectives, such as promoting competition and maximizing economic value. Many economic historians and philosophers cite spectrum auctions as a clear example of effective economic model application in policy-making (e.g. Alexandrova 2006, 2008; Alexandrova and Northcott 2009; Backhouse 2010).

Both of these examples share a common theme: they concern the interaction between economists and policymakers, whether government entities, central banks, or international organizations. However, a different form of interaction exists between economists and those who are responsible for administering justice, including courts and, to a lesser extent, juries. Armed with economic theory and econometric techniques, economic experts have entered the courtroom to offer their expertise in the pursuit of justice. One might even say that just as macroeconomists use their models to assist governments and banks in making better economic policy decisions, industrial economists employ their microeconomic models to assist courts in making more informed antitrust enforcement decisions.[4]

It is important to note, however, that such an idealized picture of how these interactions unfold in the courts does not always reflect reality. The adversarial system, typical of common law countries such as the US and UK, means that economists, like other experts, function as expert witnesses for their respective parties.

Legal epistemologist Susan Haack (2014) distinguishes between "inquiry" and "advocacy." Inquiry, associated with scientific research,

4 A potential objection to my work is that I may not have adequately considered the feedback on industrial organization ("IO" for short) resulting from economists acting as experts – a phenomenon evident in the case of macroeconometric modeling (see, for example, Acosta and Pinzón-Fuchs 2019; Boumans and Duarte 2019; Renault 2023). This objection suggests that IO models might be developed with awareness of their potential application in the legal context. I have two responses to this argument. On the one hand, it is hard to deny that many economic models were constructed in an effort to influence antitrust enforcement, including many foundational works of the Chicago approach to antitrust (see Giocoli 2022). Moreover, recent quantitative modeling techniques, such as merger simulations, have been developed for use by enforcement agencies in the context of merger litigation. On the other hand, the characteristics of the adversarial system are likely to constrain the feedback effect discussed. The judge's significant discretion in admitting scientific evidence, coupled with the fact that each trial has a different judge, poses difficulties in developing models tailored for courtroom testimony. Economists often resort to adapting pre-existing models developed in the academic context for specific legal cases (see below chapters 3 and 4). This contrasts with macroeconomists who consistently work for institutions or governments with explicit policy goals. To be sure, if there had indeed been such a feedback, the results would be dismal, considering that in more than a third of the cases economists subjected to *Daubert* scrutiny were excluded from the trial (see appendix A).

seeks to uncover the truth behind a question of interest. In contrast, advocacy aims to present the strongest case for a predetermined answer or conclusion. A genuine inquirer "is motivated to seek out all the evidence he can; to judge as fairly as possible how strong it is; in what direction it points; and how clearly; and to draw a conclusion only when he judges the he has adequate evidence to do so." (*ibid.*, p. 12). An advocate's business is a very different one: "[he] will try to seek out all the favorable evidence he can, and try to play down or explain away any evidence unfavorable to his predetermined solution." (*ivi*)

Now, one might think that expert witnesses, in contrast to "real" scientists, behave mainly as advocates for the party that has retained their services. This criticism is often encountered in the media, suggesting that expert witnesses, being financially compensated by their respective parties, are more inclined to be partisan rather than impartial and dedicated scientific investigators.[5] In fact, things are more complicated than that.

American legal scholar and federal appellate judge Richard Posner (1999) explained that there are also incentives for expert witnesses to act as scientific inquirers in the sense describe by Haack.[6] Posner offers an initial reason:

> Because most expert witnesses, including most economic expert witnesses, are repeated players (unlike most lay witnesses), they have, like other potentially disloyal agents, a financial interest in creating and preserving a reputation for being honest and competent. (Posner 1999, pp. 93–94)

While maintaining a positive reputation is indeed an incentive in expert testimony, in contentious legal domains like antitrust, the impact on reputation extends beyond mere "honesty" and "competence." Economists providing testimony in antitrust cases may strategically shape their reputation as either "pro-businesses" or "pro-enforcement." A reputation as a pro-business economist can lead to more expert testimonies for corporations facing antitrust allegations. In contrast, building a reputation as a pro-enforcement economist opens up future job prospects within government agencies responsible for antitrust enforcement, such as the Federal Trade Commission. To establish and safeguard a reputation as either pro-business or pro-enforce-

5 See, for instance, the following article on ProPublica.

6 Posner's arguments about existing incentives for expert witnesses to act as scientific inquirers are primarily theoretical. Studies in the field of psychology, particularly those conducted by Dror and Hampikian (2011) and Dror (2016), have explored a similar theme from a different angle – namely, whether the existence of cognitive bias impacts the impartiality and reliability of expert decision-making. Their findings reveal that forensic experts analyzing DNA and fingerprint evidence from crime scenes can be biased by irrelevant contextual information, such as knowledge about a suspect's ethnicity or criminal record.

ment, economists may actively engage in advocacy behaviors, which could involve downplaying or disregarding conflicting evidence.

Posner gives us a second reason to believe that expert witnesses can indeed function as scientific inquirers:

> An expert witness who has a record of academic publication will be "kept honest" by the fact that any attempt to repudiate his academic work on the stand will invite devastating cross-examination. (*ibid.*, p. 94)

The rationale behind this is straightforward: if an expert witness has published an academic paper on topic A claiming that p is true, she has a strong incentive not to state during the expert's testimony on topic A that p is false. To do so would almost certainly trigger a harsh cross-examination, exposing inconsistencies in the expert witness's beliefs and severely undermining their overall reliability. Posner's insight proves accurate, as there are instances where lawyers have sought to demonstrate that expert witnesses repudiated their own academic work. In the landmark IBM lawsuit, plaintiff's attorneys repeatedly attempted to show the inconsistencies of the expert report provided by MIT economist Franklin Fisher with his previous academic publications.[7]

Nonetheless, the binding power of previous academic publications has a big limitation: it is applicable exclusively to expert witnesses who have had or currently maintain an academic career. While there is a growing number of academic economists who offer expert testimony in legal proceedings, a substantial portion of expert witnesses are drawn from consulting firms.[8]

Finally, Posner identified a third reason why expert witnesses may be motivated to act as scientific inquirer:

> An expert witness' evidence is inadmissible if it does not satisfy the methodological standards in the expert's field – something that is easier for judge to determine than whether the analysis is correct. This rule acts as a screen against "junk science." (*ibid.*, pp. 94–95)

This third motive introduces key elements that will form the main topics of this dissertation. To begin, Posner identifies a crucial function that courts must fulfill when confronted with expert testimony: the so-called "gatekeeping role." By exercising their gatekeeping role,

7 Franklin M. Fisher papers, Duke University David M. Rubenstein Rare Book & Manuscript Library, box CW12.

8 Incidentally, it is worth noting that eminent economists and antitrust experts often hold top positions in economic consulting agencies that specialize in providing expert testimony for legal disputes. For instance, Richard Posner was one of the founders of *LexEcon*, and over half a century ago, Franklin Fisher established the "Antitrust and Competition Practice" at *Charles River Associates*.

courts possess the authority to exclude expert testimony from the trial: much like how an eyewitness can be deemed unreliable and excluded from the trial, an expert witness can also face exclusion if her testimony is deemed unreliable.

Second, he underscores that such a gatekeeping role must be exercised by verifying whether the expert's testimony aligns with the methodological standards of her field. In essence, courts are required to scrutinize the methods employed to obtain the results presented in expert testimony.

Third, he emphasizes the existence of a guiding rule that empowers judges with gatekeeping authority, directing their efforts to keep unreliable or unsound scientific evidence out of the trial. This rule, as referenced by Posner, is the standard articulated by the Supreme Court in the case of *Daubert v. Merrill Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993). The *Daubert* ruling, subsequently refined in two cases – *General Electric v. Joiner*[9] in 1997 and *Kuhmo Tire Co. v. Carmichael*[10] in 1999 – insisted that trial court judges adopt a gatekeeping role to "ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable." (*Daubert*, p. 589)[11]

Posner's observation provides the backdrop for the subject of my inquiry. Economists strive to bolster their party's position through expert testimony in legal cases. At the same time, federal judges, acting as gatekeepers, wield the power to decide whether to admit or exclude expert testimony. Their gatekeeping decisions rely on the *Daubert* standard, a rule derived from a Supreme Court decision. As said above, this study investigates such a dynamic interplay between economists and judicial institutions, specifically concentrating on the realm of antitrust law.

I chose to focus on antitrust law for several reasons. First, in the domain of antitrust law, the role of economists who testify is distinct from many other areas of law. They are not mere damage calculators or statisticians; instead, they are called upon to provide a critical contribution rooted in theoretical interpretation, drawing from economic theory (mainly, industrial economics) and the specific facts of the case at hand. Their task is to bridge the gap between economic principles and the real-world circumstances of each case, offering a comprehensive analysis that is both theoretically sound and factually relevant. The significance of economics also explains the substantial number of *Daubert* decisions in antiturst cases made by the federal courts each year.

Economic theory provides the essential framework for many of the key concepts expressed in the two main statutes of American antitrust

9 522 U.S. 136, 136-37 (1997).

10 526 U.S. 137, 138-39 (1999).

11 The interpretation of the *Daubert* ruling is still a subject of ongoing debate, so much so that a reform is currently in the approval process. For further details, see below, chapter 4.

law: the Sherman Act (1890) and the Clayton Act (1914).[12] Let us consider the two foundational concepts of antitrust law: competition and market. Economic theory provides content to both these concepts and the way we attribute meaning to these concepts changes as economic theory evolves. For example, what constitutes a competitive market? Some might define it by examining structural aspects of the market, such as the number of existing firms or their market shares. Others might focus on the presence of significant barriers to entry or how easily consumers can switch to alternative products in response to price increases. Throughout the history of economic thought, the relative importance given to these factors has fluctuated significantly. Therefore, the application of antitrust laws is intricately intertwined with the theoretical shifts in economic thought.

Two historical examples vividly illustrate the profound influence of economic theory on antitrust enforcement.[13] First, we can trace the impact of scholars affiliated with the "Chicago" and "Harvard" schools on court decisions. In the late 1970s, a shift occurred in how certain business practices, once considered antitrust violations, were reinterpreted as pro-competitive behavior. A notable illustration is predatory pricing, where a firm deliberately sets low prices to drive competitors out of the market. Under the influence of the writings of Areeda and Turner (1975) and Bork (1978), the Supreme Court increasingly dismissed predatory pricing as a practice of concern for antitrust law due its alleged unprofitability. In the 1993 *Brooke* decision, the Supreme Court established legal standards and limitations for predatory pricing claims, making it more challenging for US antitrust authorities to obtain convictions for predatory behavior.[14]

Another clear demonstration of the dynamic relationship between economic theory and antitrust enforcement is seen in the evolution of Horizontal Merger Guidelines (HMG) since the 1960s. Although these guidelines are not legally binding, they serve as essential reference points for judges when evaluating whether to allow or deny business mergers. Over time, they have adapted to align with the evolving economic discourse on the competitive consequences of mergers. For instance, the traditional emphasis on market shares as indicators of the anticompetitive nature of mergers has waned in favor of a more

12 The full texts of the Sherman Act and Clayton act are available here and here.

13 The literature exploring the historical links between economic thought and antitrust is quite extensive. Medema (2011) offers an overview of the Chicago law and economics movement. For an analysis on the role of Posner in the law and economics movement, see Marciano (2023) and Harnay and Marciano (2009). Additionally, see Kirat and Marty (2021), where the authors elaborate on how early members of the Chicago school and institutionalist economists came to endorse antitrust law enforcement in the second half of the 1930s.

14 *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993). For a history of predatory pricing in law and economics, see Giocoli (2014).

nuanced evaluation that includes alternative metrics such as diversion rations or upward pricing pressure.[15]

A second reason to favor antitrust over other legal areas is its impact on the business world, consumer welfare, and market dynamics. The celebrated paper "The Rise of Market Power and the Macroeconomic Implications" (De Loecker, Eeckhout, and Unger 2020) has documented empirically a thirty-year trend of increasing market power among large corporations.[16] One cause of such increase has been found in the decrease of antitrust enforcement in the last decades. As researchers have demonstrated, since 1990 the amount of resources devoted to federal antitrust agencies has declined despite a simultaneous increase in merger activity and in the complexity of economic reasoning required to win in the courtroom (Kwoka 2020; Lancieri, Posner, and Zingales 2022; Wollmann 2019).[17]

Furthermore, the past decade has witnessed a series of high-profile antitrust cases that have significantly reshaped the perception of both public opinion and experts on the relationship between the state and the market. *Ohio v. American Express Co.* reached the Supreme Court and marked a significant moment in antitrust cases by focusing on the unique characteristics of two-sided platforms.[18] Another noteworthy antitrust case, *United States v. AT&T Inc.*, held a distinct place in legal history because it was the first vertical merger case to be litigated in a federal court in decades.[19] Incidentally, both of these significant cases resulted in government defeats.

As I write these pages, a Washington D.C. judge is presiding over a potentially landmark antitrust case brought by the DOJ against Google. This ongoing case underscores the enduring relevance and significance of antitrust law in contemporary society. Hence, it becomes particularly intriguing to explore antitrust enforcement from a less conventional perspective – that of methodological analysis concerning the role of economic experts and the economic theories they employ in courtroom.

The debate surrounding how to deal with highly concentrated markets, particularly in the digital sector, has become a prominent topic in both US and European political arenas.[20] Within this discourse,

15 See Hovenkamp and Shapiro (2018) for an analysis of how HMG have changed over time.

16 These results are currently a matter of debate in the academic field. For an example of such a debate, see the 2019 symposium of the *Journal of Economic Perspectives* (Vol. 33, No. 3, Summer 2019) where issues in the measurement of markups and the implications of market power for the labour market are discussed at length.

17 I delved into these topics in greater detail in Paseyro Mayol and Peruzzi (2023).

18 138 S. Ct. 2274, 2018.

19 310 F. Supp. 3d 161, DC District Court, June 2018.

20 Two recently published books provide a thorough examination of the competition dynamics surrounding Big Tech. The first, *The Profit Paradox* by economist Jan Eeckhout (2021), presents an informal yet comprehensive summary of economics research spanning the past decade, delving into the market power wielded by major corporations. In contrast, the second, *Big Tech and the Digital Economy* authored by Law Pro-

two extreme positions have emerged.[21] First, there are those who argue that large corporations must be dismantled due to the excessive market power they wield, which, in their view, has granted them undue political influence over states, workers, and small businesses. Advocates of this perspective are often referred to as Neo-Brandeisians, drawing inspiration from the ideas of Louis Brandeis, a former Associate Justice of the Supreme Court from 1916 to 1939.

Advocates of the Chicago School are opposed to Neo-Brandeisians by focusing on the self-corrective character of the market system and on the risks associated with strict state regulation and antitrust enforcement. As markets tend to be competitive if not disturbed, the Chicagoans argue, wrongly convicting innocent firms has more adverse social consequences than failing to punish the guilty ones.

These two approaches not only diverge in their specific viewpoints on various aspects of antitrust law, but also in their perceptions of the role economics should play in interpreting and applying antitrust laws. Neo-Brandeisians tend to minimize the role of economics and advocate that antitrust should be guided by non-economic principles, such as the protection of local small businesses. In contrast, the Chicago School has traditionally emphasized that antitrust law should prioritize economic efficiency and asserts that sound economic theory should determine whether a particular behavior promotes efficiency. The fundamental message of this work is that both of these approaches are flawed because they evade the main methodological question: What kind of knowledge can we derive from the use of economic theory?

In the upcoming chapters, we will delve into how economic models are constructed, enhanced to reflect real-world complexities, and applied to specific cases using real-world data and econometric techniques. The resulting picture may vary in shades of gray or rose, depending on one's perspective. On one hand, economic theory emerges as an indispensable tool for understanding which business practices may lead to anticompetitive effects and for quantifying those effects. However, we must remain cognizant of the limits in the knowledge we can derive from economics.

Applying economic models to particular circumstances is a challenging task, especially in the context of legal proceedings, and it demands a high level of expertise from economists. Judges themselves grapple with the task of evaluating the admissibility of economic arguments applying the *Daubert* standard. The fact that dozens of antitrust courts every year question the reliability of economic theory and the conditions for its proper use should serve as clear motivation for the research presented here.

fessor Nicholas Petit (2020), offers a different angle by studying the various forms of competition challenging apparent monopolies, while also shedding light on the limits of traditional economic analysis for antitrust policy

21 I draw this two-camp distinction from Shapiro (2021a).

The dissertation begins with chapter 2, which defends the viability of de-idealization strategies in economic modeling against recent criticism. De-idealization occurs when an idealized assumption of a theoretical model is replaced with a more realistic one. Recently, some scholars have raised objections against the possibility or fruitfulness of deidealizing economic models, suggesting that economists do not employ this kind of strategy. We present a detailed case study from the theory of industrial organization, discussing three different models, two of which can be construed as de-idealized versions of the first (the so-called Bertrand model of oligopoly). We conclude that recent pessimism about de-idealization in economics is largely unfounded, and that de-idealization strategies are not only possible but also widely employed in economics.

Economists are often called upon as expert witnesses by the parties involved in antitrust litigation. One challenge they may face in US federal courts is compliance with the *Daubert* standard of admissibility of expert testimony. The interplay between model applicability and the *Daubert* standard is analysed in chapter 3, suggesting the importance of distinguishing between weak applicability claims, those that state that a model's critical assumptions are shared by the target, and strong applicability claims, those that connect empirical models and quantitative market features. Recent antitrust cases in which expert testimonies based on economic models have been assessed following the *Daubert* standard are examined using this framework. Some normative implications are drawn concerning how to improve judges' assessment of model-based arguments.

The quantitative impact of *Daubert* challenges on antitrust litigation is evident in the available data. In the Appendix A at the end of the dissertation, I present a simple empirical analysis illustrating the influence of the *Daubert* standard in antitrust litigation. Despite certain limitations concerning data quality, the Appendix reveals that antitrust economists encounter more frequent gatekeeping challenges in antitrust litigation compared to other legal domains. Furthermore, it provides descriptive statistics on *Daubert* challenges faced by economists in antitrust cases.

Finally, chapter 4 takes us down a skeptical slope, demonstrating the existence of the "gatekeeper dilemma." First, we argue that the main methodological choice that federal courts make during a *Daubert* hearing is whether to assess the reliability in the abstract of the methods used by the expert witness (reliability questions) or to check whether the application of those methods was correct (applicability questions). Second, we show that both choices, have undesirable results. If courts merely ask reliability questions, then they misinterpret Rule 702 and are likely to admit expert testimony at trial based on misapplications of economic theories and econometric techniques. On the contrary, if courts also pose applicability questions, then the

lack of clear and widely accepted rules of applicability in economics makes it impossible to answer such questions.

The thesis concludes with a short chapter 5, in which I lay the groundwork for future research. A potential extension of this work, I argue, could focus on the European legal system and explore the impact of economic theory on judicial decisions beyond *Daubert*.

2

DEFENDING DE-IDEALIZATION IN ECONOMIC MODELING: A CASE STUDY

No mere fact ever was a match in economics for a consistent theory.

— Paul Milgrom and John Roberts[1]

2.1 INTRODUCTION

Theoretical models in science, and in economics in particular, typically contain idealizations of various kinds. While this is widely acknowledged, there is much less consensus about the nature and epistemic role of those idealizations. One important view of idealizations refers to the notion of de-idealization (sometimes called "concretization"). According to such view, important episodes of theoretical change in science can be construed as instances of a process that goes from more idealized to more realistic models of phenomena (Nowak 1980; Cools, Hamminga, and Kuipers 1994; Niiniluoto 2002, 2012, 2018; Hindriks 2012).

In recent discussion on the methodology of economics, the notion of de-idealization and its role in the practice of the discipline has been strongly criticized (see, in particular, Alexandrova 2008; Alexandrova and Northcott 2009; Reiss 2012). Despite having different views of idealizations and economic modeling, such critics agree on one point: that de-idealization strategies are actually not used in economic modeling, for the good reason that they are either unfeasible or useless. In short, thinking in terms of idealizations and de-idealizations is not a useful or viable approach to understand scientific change and progress, at least as far as theoretical models in economics are concerned.

This chapter aims at rebutting this criticism; in particular, we put forward two main claims. First, we argue that the notion of de-idealization is sound and viable as a way of analyzing the relations between different but connected economic models. Second, we show that de-idealization strategies are actually employed by working economists, pointing to a relevant case study concerning the theory of industrial organization. We conclude that the idea of de-idealization provides both a fruitful theoretical perspective on economic modeling, and a way of reconstructing the actual historical development of the discipline.

1 "Informational asymmetries, strategic behaviour, and industrial organization", Milgrom and Roberts (1987, p. 185).

We proceed as follows. In Section 2.2, we introduce the strategy of idealizations and de-idealizations, and present some criticism raised against it within recent discussion about economic modeling. In Section 2.3, we summarize some classical discussion of idealizations and de-idealizations and suggest how this notions can be applied to the case of economic models. To better illustrate our thesis, in Section 2.4 we present a case study from the theory of industrial organization. More specifically, we discuss three different models of oligopoly markets (the standard Bertrand model, the Bertrand model with differentiated goods, and the Varian model) and highlight their conceptual relationships.

In Section 2.5, we re-evaluate the criticism advanced against the viability of de-idealizations in economics in the light of our case-study. After proposing a working definition of de-idealized models, we argue that critics tend to underestimate the importance of de-idealization techniques in economic modeling. Indeed, a striking contrast exists between the claim that de-idealizations are unfeasible and the economists' effort to relax idealized assumptions to get models with better explanatory and predictive power. We conclude that such criticism is ultimately ill-founded, and that de-idealization strategies are not only possible but also actually employed in economics.

Section 2.6 discusses the epistemic benefits of de-idealization in economics. Finally, in section 2.7 we summarize our discussion and provide a tentative assessment of its implications for the ongoing discussion of realism, progress, and related issues in economic methodology.

2.2 ARE THERE DE-IDEALIZATIONS IN ECONOMICS?

Both scientists and philosophers widely agree that idealizations are ubiquitous and unavoidable in science. As Potochnik (2017, p. 2) puts it:

> Most basically, a science practiced by limited human beings in a complex world results in widespread idealization. Idealizations are assumptions made without regard for whether they are true, generally with full knowledge that they are false. Classic examples are the assumption of a frictionless plane in physics and the assumption of perfectly rational agents in economics. Despite their falsity, idealizations appear in almost every scientific project and product.

Finding examples of idealized theories and models in science is indeed not difficult. In this chapter, we shall focus on economic theory (and the theory of industrial organization in particular). Even within this field, examples of idealized assumption abound: from perfect ra-

tionality (as Potochnik notes in the above quotation) to perfect mobility, zero transaction costs, perfectly divisible goods, full employment, and so on. Idealizations are indeed typical and widespread in economics, and virtually all economic models contain assumptions that are idealized to at least some degree (Mäki 1992, p. 324).

Given that idealizations are central and ubiquitous to most scientific disciplines, it is no surprise that philosophers of science devoted much attention to this notion. Different scholars studying idealizations in science (and related notions like abstraction or isolation) have proposed different ways of classifying and understanding idealized assumptions, models, and theories (Musgrave 1981; McMullin 1985; Cartwright 1989; Mäki 2000, 2012, 2020; Hindriks 2006; Weisberg 2007, 2013; Wimsatt 2007; Levy 2018).[2] Such proposals resulted in an interesting debate on the nature of idealized assumptions and on the role they play, especially in theoretical models. In particular, philosophers have acknowledged that scientists may employ idealized assumptions for a variety of purposes including isolating the essence of the problem (Nowak 1980), isolating causal mechanisms (Mäki 1992, 1994), improving mathematical tractability (Alexandrova 2006; Knuuttila 2009), use as a benchmark (Mäki 2020), and others.[3]

Although the debate is not settled, and there is no unified view of the nature and role of idealizations in science, it seems fair to say that most philosophers would agree on a common understanding of what idealized models and theories are (e.g. Mäki 2020). According to such a "minimal" definition, an idealization has at least the following two properties. First, it is a deliberate falsehood: an idealized assumption appearing (even not explicitly) in a theoretical model is a proposition which is known to be false when applied to real-world phenomena in the target domain. Second, idealized assumptions usually concern some parameter appearing in the formulation of the theory or model assuming a certain extreme value – typically zero or infinite — that is known to be unrealistic.[4] In the next sections, we shall consider some concrete examples of how idealizations work in science; for the moment, the above minimal definition will suffice to illustrate the purpose of this chapter.

2 The concept of "idealization" and its related notions – isolation, abstraction, etc. – have been discussed by many leading authors in the history of social science (from Johann von Thünen, John Stuart Mill and Max Weber to Milton Friedman and Thomas Schelling). In our chapter, however, we shall focus mainly on the recent literature in the philosophy of science about idealizations and modeling.

3 As Mäki (2020, p. 229) notes, Max Weber's concept of "ideal type" is a clear example of the use of idealization as benchmark insofar as real situations can be compared and understood in terms of deviations from ideal types. For instance, the impact of irrational behavior on economic decisions can be understood and explained only starting from the ideal behavior of the *homo œconomicus*.

4 Rice (2019) has recently argued that one cannot decompose a model into parts that are idealized and parts that are not. This does not seem to be in conflict with the minimal view just presented, which only requires that idealizations are acknowledged as falsehoods and modeled in terms of extreme values.

Interestingly, although the idea of idealization is widely studied and central to the recent philosophical debate, the companion notion of de-idealization has attracted less attention. Roughly, de-idealizing a theory or model means removing one of its idealized assumptions and replacing it with a new one that it is less idealized, i.e., more realistic in being closer to the actual phenomena. To quote again Potochnik's example, studying the effect of friction on the motion of objects moving down an inclined plane leads to developing a model that is de-idealized with respect to Galileo's original one.

At first sight, de-idealizations are as important as idealizations are in science. This was emphasized long ago by Leszek Nowak and his followers (the so-called Poznań School), who studied in detail the notion of de-idealization under the label of "concretization." According to Nowak, scientists often engage in what we can call an idealization-concretization strategy, by which highly idealized models are replaced by less idealized models.

In turn, Nowak defends the view that scientific progress can be construed as the development of increasingly more de-idealized theories and models, which are more realistic and hence superior than their predecessors.

Interestingly, some authors have explored the applications of Nowak's idealization-concretization strategy to economics. For instance, Cools, Hamminga, and Kuipers (1994) discuss the Modigliani-Miller theorem as an idealized law in Nowak's sense, and interpret subsequent modifications of the model — taking into account, for instance, corporate taxes and bankruptcy — as de-idealizing steps. In his defence of scientific realism in economics, Hindriks (2012) studies a series of theories developed for measuring the markup ratio of price over marginal cost that gradually relaxed the idealized assumptions of constant productivity growth and constant return to scale. Notably, he argues that economists aim to develop true theories using, among others, the so-called future-truth strategy, according to which the idealized assumptions contained in a model are progressively relaxed with more realistic ones. Consonant views have been advanced also by Niiniluoto (2002, 2012, 2018) who suggests that at least some of the historical changes in economic theorizing can be interpreted in this way as improvements toward models and theories which are "closer to the truth" than their predecessors. We shall come back to these ideas in the final Section 2.7 below.

Contrary to the above view, that sees de-idealization or concretization as a central notion in the methodological analysis of scientific models, some authors have recently argued that de-idealization strategies hardly work when applied to economic models and, therefore, are seldom used by economists (Alexandrova 2008; Alexandrova and Northcott 2009; Reiss 2012). For instance, Julian Reiss (2012, p. 379) bluntly argues that

> De-idealization strategies don't normally work and are therefore seldom employed [...]. Economics of information, transaction cost economics or the economics of imperfect competition do not provide de-idealized versions of the 'standard partial equilibrium model' with perfect information etc. – they're rival models.

According to Reiss, de-idealization strategies do not work because of the very peculiar way in which economists build their models. Generally speaking, the vast majority of theoretical economic models are mathematical models that describe the equilibrium of a market composed of rational, maximizing agents. This plain fact, Reiss argues, prevents the possibility of implementing de-idealization strategies where only a single assumption of a model is modified while holding everything else constant. In his own words:

> Because of the high standards of mathematical elegance, equilibrium solutions, methodological individualism and rationality economics models must comply with, it is not normally possible to tinker with individual assumptions that are deemed 'too highly idealized for the purpose at hand' while leaving others fully intact when building a new, less idealized model. (*ivi*)

A consonant position is advanced by Alexandrova (2008) and Alexandrova and Northcott (2009). Their own example concerns auction theory where, they argue, de-idealizations of unrealistic assumptions are hardly ever carried out:

> Some clearly unrealistic assumptions included perfect rationality of bidders, the number of objects for sale, absence of budget constraints on the part of bidders, etc. The technique of de-idealization could not be applied to any of these. [...] Even at an abstract level there was no one theoretical model that was supposed to represent the actual auction. (Alexandrova 2008, p. 389)

In sum, according to such authoritative voices, the concept of de-idealization has no real role to play in the analysis of how economists develop, assess, and change their models. In this chapter, we aim to show that this pessimism about de-idealization in economics is unfounded. To this purpose, in the next section we present in more detail the role of idealization as applied to economic models. We then discuss a case study taken from the theory of industrial organization and, in Section 2.5, we come back to the arguments advanced by the critics against that view.

2.3 IDEALIZATIONS AND DE-IDEALIZATIONS: SETTING THE STAGE

Before considering the case of economic modeling, let us start with a classical example of the method of idealization and de-idealization as discussed by Nowak (1980), i.e., the ideal gas law, also known as Clapeyron's law (or Clausius-Clapeyron's equation). As the name suggests, the law introduces some idealized assumptions in order to describe the behavior of a hypothetical ideal gas, which is however a good approximation of many real gases. Denoting P, V and T the pressure, volume and temperature of the gas, respectively, and R the ideal gas constant, the law states that in equilibrium:

$$PV = RT \tag{1}$$

Crucially, Clapeyron's law makes two idealized assumptions concerning the particles composing the gas: first, that all the forces of interaction a between these particles are negligible; second, that the same holds for the volume b of each particle. According to Nowak, an adequate formulation of Clapeyron's law – which makes clear that it concerns the behavior of an ideal gas composed of many randomly moving point particles that are not subject to inter-particle interactions – is hence the following:

$$\text{if } a = 0 \text{ and } b = 0 \text{ then } PV = RT \tag{2}$$

The conditional form of the law highlights its idealized nature; moreover, it also invites reasoning about its possible concretizations, corresponding to the cases where either a or b, or even both, are not negligible and hence different from zero. In the latter case, one obtains the van der Waals's law, which can be viewed as a de-idealization of Clapeyron's law:

$$(P + a/V^2)(V - b) = RT \tag{3}$$

The above law takes into account the influence of inter-molecular attractive forces (a) and the finite size of the gas particles (b), specifying what happens when both the conditions appearing in the antecedent of the original law in conditional form are false. Note that, if one puts $a = 0$ and $b = 0$ in equation 3, one obtains again the original Clapeyron's law in (1). In this sense, the idealized law is a special case of the more general (and complex) de-idealized law.

According to the Poznań School, science is littered with laws containing idealized assumptions, and scientists routinely employ the strategy of idealization and de-idealization in order to develop their models and theories. Such strategy can be described in abstract terms as follows. One starts with a law governing the behavior of some objects x and expressing a functional dependence of a quantity $F(x)$ on a finite number of other quantities $q_1(x), ..., q_n(x)$. In the simplest case, where $n = 1$, the law has the form:

$$F(x) = f_0(q(x)) \tag{4}$$

In most cases, equation (4) ignores the influence of some relevant factors $w_1(x), \ldots, w_k(x)$, which however do affect the behavior of x beyond the considered factor q. In other words, equation (4) is an idealized law which assumes that $w_1(x), \ldots, w_k(x)$ are all negligible and equal to zero (or to some suitably specified value). [5] If just one such factor w_1 is ignored, the proper form of the law is expressed as follows:

$$\text{if } w_1(x) = 0 \text{ then } F(x) = f_0(q(x)) \tag{5}$$

The conditional form makes the role of the idealized assumption concerning $w_1(x)$ explicit: the value of the magnitude w_1 attributed to the object x is assumed to be zero even though, as a matter of fact, this is not the case.[6] De-idealizing such a law amounts to removing the idealized assumption by taking into account the ignored factor w_1, thus obtaining a new, de-idealized law with the following form:

$$\text{if } w_1(x) \neq 0 \text{ then } F(x) = f_1(q(x), w_1(x)) \tag{6}$$

In principle, this process can be repeated, in a step-wise fashion, for any other factor that scientists find important to take into account, leading to more and more complex, and de-idealized, versions of the original law.

Nowak's approach, as outlined above, provides a quite abstract framework to deal with idealizations and de-idealizations as far as mathematical laws are concerned. To the best of our knowledge, a formal definition of "idealized" and "de-idealized" models comparable in rigor to Nowak's one for laws remains to be developed.[7] Still, it is not difficult to find, in the economic literature, examples of more or less idealized models that are connected to each other by relations that can be quite naturally construed as de-idealization relations in Nowak's sense. Let us briefly mention but one example, that we discuss later in detail, of how de-idealization strategies work in the case of economic models.

The theory of Industrial Organization ("IO theory" for short) is a branch of economics which deals with the strategic behavior of firms

5 To be more precise (cf. Nowak 1980, pp. 28–29), here "zero" represents the minimum value that the magnitudes w_i can meaningfully assume in the relevant context. In some cases, as with the Clapeyron's law seen above, this is actually zero. In others, this zero-value represents some relevant constant, meaning for instance that the change in the quantity w_i is zero, or that some ratio of relevant factors is constant, and so on (see Section 2.4.1 for an example). In short, in an idealized law or model, some parameter assumes an extreme value (not necessarily zero) which is unrealistic and makes some quantity or factor negligible in the relevant context.

6 Here we neglect the semantic problem concerning how to interpret Nowak's idealized laws. See Niiniluoto (2002, 2018).

7 In Section 2.5, we make an attempt to provide such a definition. See also Knuuttila and Morgan (2019) for a detailed analysis of many aspects of the notion of de-idealization.

in imperfectly competitive markets and with its implications for policy related topics such as antitrust regulation and mergers (i.e., separate firms that join together to form a single one). A large part of current IO theory focuses on oligopolies, i.e. markets with a small number of large sellers (so-called oligopolists) and a big number of consumers that demand the goods produced by those firms. One core model of IO theory is the so-called Bertrand model; it concerns a simple market containing just two firms, both producing exactly the same good at constant marginal costs. Clearly, many assumptions of this model are highly idealized: to mention but a few, perfect homogeneity of goods, perfect information among consumers, no capacity constraints, constant marginal costs, profit maximization and so on. With such highly unrealistic assumptions, one is able to prove that the equilibrium price in the market equals the marginal cost — a result which is hardly ever supported by empirical evidence and, for this reason, is known as the "Bertrand paradox."

Economists have reacted to the Bertrand paradox in various ways, often developing new models that modifies the original one by relaxing one or the other of its unrealistic assumptions. A case in point is the assumption of perfect homogeneity, which is a clear instance of an idealization, since real-world markets are characterized by a sharp degree of product differentiation in terms of physical characteristics, location, and consumers' tastes.[8] What is known in the literature as the "Bertrand model with differentiated goods" (henceforth, the "D-Bertrand model" for brevity) has been developed exactly to avoid the assumption of perfect homogeneity and to account for product differentiation. The new model keeps essentially unaltered all the assumptions of the Bertrand model, but assumes that the two firms produce two different goods (e.g., Coca-Cola and Pepsi). In this way, the D-Bertrand model can account for product differentiation and make sense of firms' market power, thus avoiding the Bertrand paradox (see section 2.4.1 for details).

The above example makes clear some general features of idealizations and de-idealization in economics that are worth noting. First, both the Bertrand and the D-Bertrand model aim at describing oligopolistic markets. Second, whereas both models are highly simplified and hence idealized, the D-Bertrand model is less idealized that the Bertrand model, because it avoids the perfect homogeneity assumptions that the latter makes. Third, and crucially, the D-Bertrand model provides a measure of product differentiation which allows one to say how diverse are the two goods on the market; when such measure is negligible, the original Bertrand paradox re-appears, and the two models basically collapse on each other. In this sense, the D-Bertrand

8 Eaton and Lipsey (1989) provide an extensive treatment of product differentiation. See also Belleflamme and Peitz (2015, Chapter 5).

model de-idealizes the standard Bertrand model in a similar fashion to how van der Waals's law de-idealizes Clapeyron's law.

As we shall better argue in Section 2.5, the notions of idealization and de-idealization, understood in a broadly Nowakian sense, provide a fruitful and viable to analyze the theory and practice of economic modeling. To corroborate our thesis, in the next section we present in some detail three different models within IO theory, exploring their conceptual relationships in the light of our previous discussion.

2.4 CASE STUDY: MODELS OF OLIGOPOLY PRICING

The French engineer Joseph Bertrand (1883) introduced his model while discussing another classical contribution to the theory of oligopoly, the Cournot duopoly model. Later, Francis Edgeworth (1925) further developed Bertrand's ideas and today the Bertrand model of oligopoly pricing is part of standard IO textbooks.

The Bertrand model describes the interaction among sellers and buyers in a highly idealized market, with the goal of predicting the price charged by sellers. It considers two firms 1 and 2 (the sellers), both of which produce the same homogeneous good at identical constant marginal costs c. The firms simultaneously set prices in order to maximize profits. On the demand side of the market, there are the firms' customers (the buyers), who always buy from the cheapest seller. The demand function faced by firm 1 is simply defined as follows, where $Q_1(p_1)$ is the quantity of the good demanded by the buyers at price p_1:

$$Q_1(p_1) = \begin{cases} Q(p_1) & \text{if } p_1 < p_2 \\ Q(p_1)/2 & \text{if } p_1 = p_2 \\ 0 & \text{if } p_1 > p_2 \end{cases} \tag{7}$$

In words, this means that, as consumers are fully responsive to changes in prices, when firm 1 charges a lower price than firm 2, the former attracts the entire demand. Conversely, if firm 1 charges a higher price than firm 2, then no consumers will buy from 1. Finally, if both firms charge the same prices, then the demand is equally split between the two firms. The demand function of firm 2 is defined in the same way, just replacing "1" with "2" in the above formula.

In this model setup, a single Nash equilibrium in pure strategies exists where both firms set prices equal to their marginal costs. This is easily seen by noting that, for all other price combinations, at least one firm has an incentive to deviate, i.e., to set a price different from c. If $p_1 > p_2 > c$, then the firm 1 can increase its profits by setting a price $p_1' \in (c, p_2)$; if $p_1 = p_2 > c$, then each firm can slightly

undercut the rival price to increase profits; if $p_1 > p_2 = c$, then firm 2 can increase its profits by increasing price above c and just below p_1. Thus, $p_1^* = p_2^* = c$ is the unique Nash equilibrium of the game (the star denoting equilibrium prices).

This result – known as the "Bertrand paradox" – shows that price competition between two firms is sufficient to equalize all prices at the level of the marginal cost. This is called a paradox as, in real markets, this does not happen. Indeed, even a casual observation of economic reality immediately suggests that firms usually set their prices above their marginal costs, i.e., they have at least some form of "market power." As a consequence, a great deal of work within industrial economics has been devoted to understanding where market power comes from and how firms manage to raise their prices above marginal costs in the face of Bertrand's result.

Researchers in IO have identified two main sources of firms' market power in oligopolistic markets. First, goods are almost never perfectly homogeneous. Second, consumers normally experience a degree of inertia whereby they are not fully susceptible to changes in price. The first aspect of real (as opposed to idealized) markets is known as "product differentiation," the second as "consumer inertia." In the remaining part of this section, we consider two well-known models – i.e., the Bertrand model with differentiated goods and the so-called Varian model – that take into account, respectively, product differentiation and consumer inertia.

In the next section, we shall argue that these two models can be construed as two different de-idealizations of the original one in the sense defined in Section 2.3.

2.4.1 *The Bertrand model with differentiated goods*

The standard Bertrand model assumes that products are perfectly homogeneous so that consumers cannot differentiate among brands or distinguish among the producers when purchasing a specific product. Economists find this assumption highly unrealistic for the majority of markets and they employ various approaches to modeling industries producing differentiated products. Here we follow the approach developed by Dixit (1979) and Singh and Vives (1984). Their idea is pretty simple: two firms produce two differentiated products facing a linear inverse demand curve from consumers that gains utility from consuming a variety of goods.[9]

9 As economists know, this is by no means the only way to modeling product differentiation. Another well-known approach is the so-called "location approach," originated by the seminal article by Hotelling (1929). Both approaches are now part of standard IO textbooks such as Shy (1995), Martin (2010), and Belleflamme and Peitz (2015).

Suppose that two firms produce differentiated brands of a soft drink – e.g. Coca-Cola and Pepsi – with linear inverse demand functions (brands' prices as functions of quantities):

$$p_1 = \alpha - \beta q_1 - \gamma q_2 \tag{8}$$

$$p_2 = \alpha - \gamma q_1 - \beta q_2 \tag{9}$$

where (p_1, q_1) and (p_2, q_2) are the price and quantity of, respectively, Coca-Cola and Pepsi, and α, β, $\gamma \in \mathbb{R}$ are parameters. Note that, contrary to what happens in the original Bertrand model, equations (8) and (9) together imply that the price for Coca-Cola depends both on the quantity of Coca-Cola and on the quantity of Pepsi, and vice-versa. Moreover, it is assumed that $\beta > 0$ (to have standard downward sloping demand curves) and that $\beta^2 > \gamma^2$. The latter assumption is crucial because it implies that the price of a good is more sensitive to a change in the quantity of that good than to a change in the quantity of the other one. In the jargon of economists, this means that the "own-price effect" dominates the "cross-price effect."

By inverting (8) and (9) we find the system of direct demand functions (quantity demanded as functions of brands' prices):

$$q_1 = a - bp_1 + dp_2 \tag{10}$$

$$q_2 = a + dp_1 - bp_2 \tag{11}$$

where

$$a = \frac{\alpha(\beta - \gamma)}{\beta^2 - \gamma^2} \tag{12}$$

$$b = \frac{\beta}{\beta^2 - \gamma^2} > 0 \tag{13}$$

$$d = \frac{\gamma}{\beta^2 - \gamma^2} > 0 \tag{14}$$

The appealing feature of this setup is that one can formally define a measure δ of product differentiation, as follows:

$$\delta = \frac{\gamma^2}{\beta^2} \tag{15}$$

This means that two products are highly differentiated when the cross-price effect tends to zero, i.e. $\gamma^2 = 0$ and, therefore, $\delta = 0$. Intuitively, consumers perceive products so differently that a change in the price of Coca-Cola has little or no influence on the demand for price of Pepsi. Conversely, the two products are perfectly homogeneous when the cross-price effect is equal to the own-price effect, i.e.,

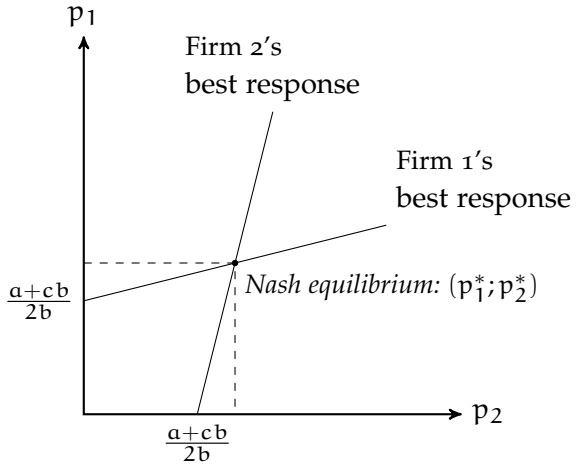


Figure 1: The D-Bertrand model: Nash equilibrium price is given by the intersection between firms' best response to each other. The best-response functions are upward sloping, meaning that if one firm raises its price, the other would respond by raising its price as well. The value $\frac{a+cb}{2b}$ is the optimal price charged by each firm when the other one sets the price equal to zero; formally, it is the intercept of the two best response functions with the axes.

$\gamma^2 = \beta^2$ and, therefore, $\delta = 1$. This means that if consumers consider the two products as almost homogeneous, a change in price of Coca-Cola will have a great influence on the demand of Pepsi. This is the mechanism that we saw at work in the standard Bertrand model: a slight difference in price causes a great shift in consumers' demand.

As for the standard Bertrand model, it is assumed that firms simultaneously set prices in order to maximize profits given identical constant marginal costs c. The unique Nash equilibrium (see Figure 1) can be found by considering the mutual best-responses of firms (i.e., the profit-maximizing price of Coca-Cola for an arbitrary price by Pepsi and vice-versa):

$$p_1 = \frac{a + dp_2 + cb}{2b} \tag{16}$$

$$p_2 = \frac{a + dp_1 + cb}{2b} \tag{17}$$

and, thus,

$$p_1^* = p_2^* = \frac{a + cb}{2b - d} = \frac{\alpha(\beta - \gamma)}{2\beta - \gamma} + \frac{c\beta}{2\beta - \gamma} \tag{18}$$

Prices and profit levels increase when products are more differentiated, i.e. when γ tends to zero. This means that product differentiation increases the market power of firms by relaxing price competition. Conversely, prices and profit levels decrease when the products

become less differentiated, i.e. when γ tends to β. By the way, this explains why Coca-Cola and Pepsi spend millions on advertising to convince consumers about the uniqueness of their products.

The D-Bertrand model is a cornerstone of modern industrial organization. First, it sheds light one important source of corporate market power. Second, it plays a prominent role in contemporary merger analysis where it is employed to predict the post-merger equilibrium prices after being calibrated using data from a given industry (Werden, Froeb, and Scheffman 2004; Einav and Levin 2010).[10] For our purposes, what is important to note is how this models compares with the original Bertrand model. In this connection, one can easily check that the Bertrand paradox – i.e., the result that $p_1^* = p_2^* = c$ – can be obtained as a special case of the new model when $\gamma = \beta$ or, equivalently, $\delta = 1$:

$$\text{if } \delta = 1 \text{ then } p_1^* = p_2^* = \frac{\alpha(\beta - \beta)}{2\beta - \beta} + \frac{c\beta}{2\beta - \beta} = \frac{c\beta}{\beta} = c \tag{19}$$

In other words, the original Bertrand model is a special case of the D-Bertrand model when the goods produced by the firms are perfectly homogeneous.

2.4.2 *The Varian model of sales*

A second source of firms' market power comes not from the intrinsic features of their products, but from exogenous elements characterizing their market environment. One of these elements is that some consumers may be more informed than others regarding the prices charged by different firms. The American economist Hal Varian (1980) famously exploited this fact in order to explain market power in a well-known article published in the *American Economic Review*.

The setup of the Varian model is analogous to that of the standard Bertrand, except for relaxing the perfect information assumption, thus taking into account that some consumers are more informed than others about the market prices. More precisely, Varian assumes that two firms sell a homogeneous product competing on prices in order to maximize profits. Firms have unlimited capacity to supply this product at a constant marginal costs c. This market is served by a price information clearinghouse (e.g., a newspaper or an Internet price comparison site). A number S of "informed" consumers consult the information clearinghouse and purchase at the lowest listed price. The remaining U "uninformed" consumers, by contrast, do not consult the clearinghouse and choose a firm at random as long as the price does not exceed v. This implies that each firm attracts $U/2$ unin-

10 In the next chapter, we will see another example of the D-Bertrand model employed during expert testimony in antitrust litigation.

formed consumers. All consumers have unit demand with a maximal willingness to pay of $v > c$.

As Varian shows, the heterogeneity, in terms of access to information, among consumers leads to different equilibrium prices than those predicted by Bertrand. More formally, although there is no equilibrium in pure strategies, a symmetric equilibrium in mixed strategies exists such that firms randomize prices. The intuition is the following: if one firm set its price deterministically (i.e. choosing a strategy with probability equal to 1), then the other firm could take advantage by setting a lower price and thus capturing all informed consumers. Thus, firms find it profitable to vary their prices over time, a phenomenon sometimes referred to as "temporal price dispersion"[11] It can be proven (we sketch such a proof in the addendum to this chapter) that such a mixed-strategy equilibrium occurs when each firm charges a price p in the interval $[p_0, v]$, where

$$p_0 = c + (v - c)\left(\frac{U/2}{U/2 + S}\right) \tag{20}$$

with cumulative probability $F(p)$:

$$F(p) = 1 - \left(\frac{(v - p)(U/2)}{(p - c)S}\right) \tag{21}$$

It is worth noting that price dispersion arises from the existence of an information clearinghouse which provides a subset of consumers with a list of prices charged by different firms in the market. Firms face a trade-off between lowering price to attract informed consumers (who would not buy at a price higher than the lowest one offered by the sellers) and keeping the price higher hoping to sell to uninformed customers. Thus, the two firms charge different prices for an identical good because of the heterogeneity in consumers' information.

Again, we are interested here in the relations between Varian's model and the standard Bertrand model. The crucial difference between the two models is the assumption, implicit in the latter one, that consumers can acquire information on prices at no cost. Varian relaxes this assumption differentiating informed and uninformed costumers. As a consequence, we can recover Bertrand's setup within Varian's model by assuming that, when information costs are zero, all costumers become informed, i.e., U is zero. In such case, the symmetric equilibrium distribution of prices is degenerate and the Bertrand paradox re-appears, with all firms pricing at marginal cost:

$$\text{if } U = 0 \text{ then } p_0 = c + (v - c) \cdot 0 = c \tag{22}$$

In short, if acquiring information is costless the Bertrand model turns out to be a special case of the Varian model.

11 Temporal price dispersion seemingly happens in a wide range of contexts, from sales in retail markets to life insurance industry and Internet price comparison websites. See Baye and Morgan (2001) and Baye, Morgan, and Scholten (2006).

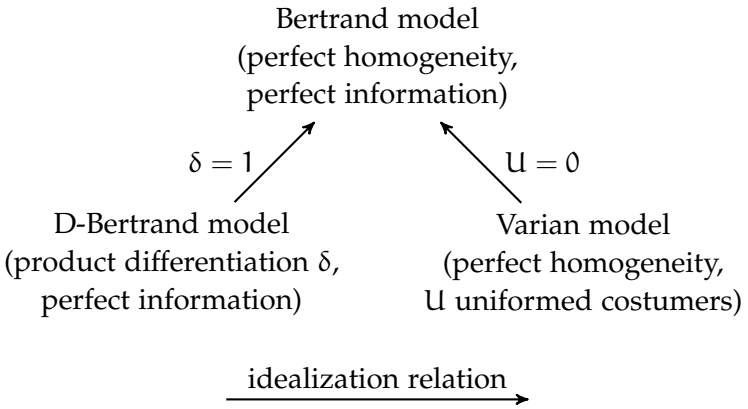


Figure 2: The Bertrand model with differentiated goods (D-Bertrand) and the Varian model as two different de-idealizations of the Bertrand model.

2.5 DEFENDING DE-IDEALIZATION IN ECONOMIC MODELING

In the foregoing section, we considered three different models of oligopoly markets in order to show how de-idealization strategies apply to the practice of economic modeling. We began with the standard Bertrand model, which assumes, among others idealizations, both perfect homogeneity of goods and perfect information of consumers. We then considered two different models, each relaxing exactly one of these two idealizing assumptions. The D-Bertrand model accounts for product differentiation, thus abandoning the perfect homogeneity assumption. The Varian model introduces information costs and uninformed customers, thus abandoning the perfect information assumption. As we argued, both the D-Bertrand and the Varian models can be clearly construed as two different de-idealizations of the highly idealized Bertrand model. This is shown by the fact that if one starts from one of the de-idealized models and re-introduce the corresponding idealizing assumption (by assigning to the relevant parameter an adequate extreme value), one comes back to the original Bertrand model. These processes are synthetically represented in Figure 2, where arrows denote the introduction of idealized assumptions.

As an upshot of our discussion so far, we can introduce the following tentative definition of de-idealized economic models. We shall say that a (mathematical) model M_2 is a de-idealization (or a de-idealized version) of another model M_1 when:

(i) both M_1 and M_2 aim at modeling the same phenomenon or domain T, what we call their "target";

(ii) with reference to T, M_1 contains at least one idealized assumption which is not contained in M_2; in this sense, M_2 is more "T-realistic" than M_1;

(iii) M_1 is a special case of M_2, i.e., M_1 can be obtained from M_2 by assigning some specific value to some parameter of M_2.

It should be clear how the above conditions apply to our case-study. With reference, for instance, to the D-Bertrand model one can see that: (i) the Bertrand and the D-Bertrand model aim at describing the same target T, i.e., a simple market where two firms set prices facing the demand from many consumers; (ii) the D-Bertrand model relaxes the perfect homogeneity assumption of the Bertrand model and accounts for product differentiation, thus being more realistic with respect to T; (iii) when the degree of product differentiation tends to zero (i.e., measure δ tends to 1), the standard Bertrand can be obtained as a special case of the D-Bertrand model. In short, the D-Bertrand model can be construed as a de-idealized version of the standard Bertrand model.

Similar considerations can be repeated as far as the relations between the Varian and the Bertrand models are concerned. Again, the two models share the same target T as above, and the Varian model is more T-realistic than the other, for it relaxes the assumption of perfect information of customers; moreover, when such assumption is re-introduced (i.e., when the number U of uninformed consumers is 0), the Bertrand model can be recovered as a special, idealized case of the Varian model.

The definition proposed above is surely incomplete and in need of improvement; still, we believe it captures, as our case-study shows, some important features of the actual practice of economic modeling. Moreover, it allows us to put in perspective the criticism advanced against the viability of de-idealization strategies in economics that we presented in Section 2.2. For instance, Reiss's skepticism about de-idealizations essentially boils down to the claim that less idealized models are not de-idealized versions of more idealized ones, but they are plainly "rival models" (Reiss 2012, 379). According to Reiss (*ibid.*),

> it is not normally possible to tinker with individual assumptions that are deemed 'too highly idealized for the purpose at hand' while leaving others fully intact when building a new, less idealized model [...] when a factor is deemed too important to be ignored, the framework is changed altogether.

This may well be true for some of the examples Reiss has in mind, but this is not sufficient to conclude that "de-idealization strategies don't normally work and are therefore seldom employed" in economics (*ivi*). Indeed, we believe that the previous discussion of our case-study shows that exactly the opposite is true.

Moreover, adopting Reiss's view would lead us to consider as rivals a whole range of models that economists consider to be extensions or special cases of other models. As a matter of fact, economists do not

portray Bertrand's model with differentiated goods as a rival to the standard Bertrand but rather as one of its extensions. Even a quick glance at the major IO textbooks supports our observation; to quote but two examples:

> We analyse here several models of price competition. We start with the standard Bertrand (1883) model where products are homogeneous. Then, we extend the model in two directions: first, we assume that firms have private information about their marginal costs of production; second, we consider differentiated products. (Belleflamme and Peitz 2015, p. 45)

> Naturally, the analysis turns out to be much less extreme [...] if the stringent and somewhat unrealistic assumption of good homogeneity is relaxed. (Vega-Redondo 2003, p. 81)

In sum, we believe that only a careful analysis of specific case-studies can tell us whether, and to what extent, economists do employ de-idealization strategies in developing their models, as we suggest they do.

In this connection, a couple of more general remarks are in order. First, a de-idealized model can in turn be highly idealized and be subject to further de-idealizations. For instance, the Varian model assumes, besides the perfect homogeneity of goods, that it is costless for firms to list prices on the clearinghouse. However, real-world firms are normally charged a certain amount to advertise their prices on newspapers or websites. For this reason, new models have been built that relax this idealized assumption and allow firms to decide whether to list their price at the clearinghouse.[12] The original Varian model can then be obtained in the limiting case when advertising costs tend to zero (Baye, Morgan, and Scholten 2006, p. 352). This is another case of the de-idealization strategy at work, consistent with the idea defended here that the de-idealization process occurs in a progressively step-wise fashion.

This is the idea that Reiss seems to critique when, in his attack on realism in economics, he writes:

> A final defence of realism in the light of the fact that all models are false is the Hegelian one of regarding models not individually but as a sequence progressing towards perfection. Individual models may well be false but error is eliminated progressively through de-idealization, deisolation and the like. (Reiss 2012, p. 379)

12 See Baye and Morgan (2001), Baye, Morgan, and Scholten (2006), and Shelegia and Wilson (2021). These authors show that equilibrium price dispersion arises – provided that advertising costs are not too large that firms refuse to list prices at the clearinghouse.

As the epithet "Hegelian" suggests, however, this is too simplistic an understanding of how the de-idealization strategy works. As our case-study clearly shows, the de-idealization process does not need to be thought as a linear, indefinite sequence of models, each less idealized than the previous one, tending, so to speak, to some Hegelian, entirely de-idealized, final model. In fact, even a quick glance at Figure 2 shows that this process may well evolve in a tree-like fashion, with multiple de-idealized models obtained from the same model, which may be quite unrelated to each other. Indeed, in our example, it is not clear in what sense the D-Bertrand and Varian model are comparable except that they are both descendant of the standard Bertrand model. To recall, while the D-Bertrand model relaxes the assumption of product homogeneity but not that of perfect information, the opposite is true for the Varian model. These two models gave rise to two rather different strands of economic literature and their convergence toward a single, "final" model where both idealized assumptions are relaxed is at least uncertain, if not unlikely. Such convergence toward a fully de-idealized model may well happen, but it is not necessarily the case; only further theoretical and empirical research, rather than general philosophical claims, may decide the issue. In any case, such convergence is not necessary to support the view that de-idealization strategies are used by economists. Indeed, our previous analysis is not affected by whether the tree structure depicted in Figure 2 converges to a single de-idealized model or diverges in a sort of bush-like family of models.

Finally, let us consider the critique of de-idealization advanced by Alexandrova (2008) and Alexandrova and Northcott (2009). These authors construe de-idealization as a strategy to test the robustness of the model, i.e., to investigate whether the conclusions of a model still hold if we change some of its theoretical assumptions. More specifically, de-idealizing a model aims to find the minimal set of assumptions such that a characteristic result holds:

> The point of de-idealization is to check that the relationship observed in the original model [...] would still hold once some of the assumptions of the original model were no longer satisfied. (Alexandrova 2008, p. 388)

While they acknowledge the usefulness of this strategy — when "the de-idealization process is successful and the model's result applies" (Alexandrova and Northcott 2009, p. 311) —, they also believe that "at least as far as economics is concerned, such happy cases may be hard to come by" (*ivi*). In particular, as already recalled in Section 2.2, they discuss a case-study from auction theory to argue that de-idealizations of unrealistic assumptions are hardly ever carried out.

In this regard, two comments are in order. First, even taken for granted that de-idealization is unfeasible in auction theory, this does

not mean that it cannot be useful in other fields of economics. As Alexandrova (2008, p. 389) herself frankly admits, the fact that de-idealization strategies are seldom used in auction theory "does not discredit the technique of de-idealization – we do know how to de-idealize some albeit not all assumptions". Second, and contrary to what Alexandrova and Northcott seem to believe, we maintain that a relevant difference exists between the epistemic goal of robustness analysis and that of de-idealization techniques. Robustness analysis is performed to discover the minimal set of assumptions which is essential for deriving a given result within the model (Kuorikoski, Lehtinen, and Marchionni 2010). De-idealization, on the other hand, amounts to replacing a false, unrealistic assumption with a more realistic one in order to obtain a better model from an explanatory and predictive point of view. This distinction, we maintain, is crucial. In fact, as our case-study shows, when one de-idealizes a model, one typically finds new and different results from those that could be derived from the original model. For instance, price dispersion among firms is a result derivable in the Varian model, but not in the Bertrand model. This may be problematic from the point of view of robustness analysis but not from that of the de-idealization strategy. In short, although robustness analysis aims at deriving the *same* results from a different set of assumptions, the whole point of de-idealization is instead deriving *different*, more realistic results from it.

2.6 EPISTEMIC BENEFITS OF DE-IDEALIZATION IN ECONOMICS

This section delves into what could be considered a central question for our analysis: the epistemic benefits of de-idealization for economists. Do de-idealization strategies offer additional knowledge to economists? If de-idealizing models does not yield any discernible benefits, what motivates economists to engage in such practices?[13]

To start, our aim in the chapter is to rationalize what appears to be a common concern among practicing economists, often expressed through various phrases like "relax assumptions", "augment mod-

13 This section is absent from the published version of this chapter (cf. Peruzzi and Cevolani 2022). Our published article prompted some responses from philosophers of economics, resulting in two main objections. The first asserts that our utilization of the notion of de-idealization is overly narrow and detached from scientific practice (Knuuttila and Morgan 2019). The second objection posits that advocating for the importance of de-idealization commits us to believing in the existence of a so-called *perfect* or *complete* model – a fully realistic model of a target system which is, by definition, in no way distorted (Finestone 2022). Although we contend that both objections overlook the essence of our argument, we postpone our response for another occasion. Indeed, we view the inquiry into the epistemic benefits derived by de-idealization as a preliminary aspect of any philosophical discussion.

els", "extend models", and so forth.[14] The immediate question that arises is: why do economists engage in these modeling practices? After all, de-idealizing a theoretical model inevitably leads to increased complexity, as evident in the examples we have examined. For instance, the Bertrand model is simpler in terms of the game structure, mathematical formulation, and the number of parameters involved in equilibrium compared to the D-Bertrand and Varian models. If this increase in complexity does not yield corresponding epistemic benefits, then the purpose of creating such models becomes unclear.

In response to this, we believe that employing de-idealization strategies in economic modeling yields clear epistemic benefits. One of the most significant advantages is that de-idealized versions of the original model typically have a broader range of applicability to real-world situations. Although the Bertrand model applies only to markets where products are homogeneous and information is perfectly distributed among consumers – a scenario admittedly found in very few markets – its de-idealized versions are applicable to a much wider range of markets where products are differentiated or acquiring information is costly. Therefore, a typical result of de-idealization is an expansion in the scope of possible applications for theoretical models.

That being said, two caveats surely exist, as we explicitly acknowledged in our discussion. First, de-idealized models are still highly idealized and thus have a limited scope of applicability. Although the D-Bertrand model is applicable to more markets than the original Bertrand model, its applicability is nevertheless limited by the presence of rather stringent assumptions about cost structure or consumer information.[15] Second, it may not be worthwhile to continue de-idealizing models indefinitely. Scientists aim to build models that are simple enough to be tractable and offer stylized representations of target phenomena. Increasing realism, therefore, yields decreasing marginal returns, so to speak: a model must be realistic enough to be reasonably applied to the intended target, but not too realistic as to impede its applicability due to its complexity.[16]

A second advantage of employing de-idealization strategies arises from establishing non-trivial connections among theoretical models. Formally proving that a model (or, more precisely, its main result) can be derived as a special case of a de-idealized model is far from trivial and leads to significant advancements in scientific knowledge. A typical scenario occurs when a previous model result is demonstrated

14 The reader can appreciate this concern by opening any standard economics textbook. Examples from areas of economics other than industrial organization are provided in the next section.

15 On the applicability of the D-Bertrand model in judicial settings, see the next chapters.

16 The trade-off between realism of models and their complexity will play a central role in our analysis of the use of economics in courtroom in the next chapters.

to be based on certain idealizations. For example, by explicitly modeling product differentiation in the D-Bertrand model, economists realize the extent to which the Bertrand paradox was founded on idealizations about product homogeneity. De-idealization strategies thus provide a retrospective rationalization of the limitations of previous modeling results. In this process, economists effectively gain new knowledge: by embedding an old result into a new model, they learn that a result previously taken for granted was actually contingent on certain idealized assumptions.[17]

To expand the scope of applicability of theoretical models and understand the limits of previous modeling results are just two reasons why economists might choose to pursue de-idealization. While additional research is needed to provide a more robust defense of de-idealization as outlined in this chapter, we believe that these insights contribute to demonstrating the importance of such a notion in studying the methodology of contemporary economics.

2.7 CONCLUDING REMARKS AND FUTURE WORK

The main claims advanced in this chapter can be summarized as follows: first, the notion of de-idealized (economic) models is sound and viable as a way of analyzing the relations between different but connected models; second, economists routinely employ de-idealization techniques in order to develop better models; third, the pessimistic outlook of some critics about such techniques is either unfounded or misplaced.

We are aware that a single case study is not enough to substantiate a fully-fledged methodological claim. On the other hand, we believe that our conceptual scheme can be adapted to many other economic contexts both within and outside IO theory. Let us consider, for instance, the case of macroeconomic theory. In this area, one finds both "classical" business cycle models developed in the 1980s and more recent models commonly adopted in contemporary macroeconomics — the so-called dynamic stochastic general equilibrium (DSGE) models. Both kinds of models share some crucial idealized assumptions: for instance, intertemporal utility maximization and the rational expectations hypothesis; however, the DSGE models de-idealize classical ones by taking into account imperfect competition, nominal rigidi-

17 Similarly, Lehtinen (2021, 2022) argues that epistemically beneficial generalizations increase the expressive power of a model. We contend that the same benefits arise in our cases. The Bertrand model lacks the expressive power to address certain aspects of information and product differentiation. For instance, the measure product of differentiation (that is, δ) only appears in the de-idealized D-Bertrand model. In this sense, the D-Bertrand model has a broader vocabulary and generalizes the original model. It is only when examining the de-idealized model that one can truly "understand" what the first one assumed, or rather, did not express at all. This perspective provides a compelling justification for the use of de-idealizations in science.

ties, and monetary policy shocks (cf. Romer 2019). For this reason, Kuorikoski and Lehtinen (2018, p. 254) recently argued that model development in macroeconomics proceeds by "vertical expansion in that new modifications and 'improvements' are added to the generally accepted but in itself simplified and idealized core model to improve its realisticness and empirical accuracy." Another compelling example would concern behavioral economics models and their possible interpretation as de-idealized versions, in the sense described here, of neoclassical models. For instance, the model of other-regarding preferences by Fehr and Schmidt (1999) or the hyperbolic discounting function popularized by Laibson (1997) are interesting case studies to which our arguments are arguably applicable. Indeed, some leading behavioral economists seem to endorse the view that behavioral models could be seen as de-idealizations of neoclassical models. For instance, Sanjit Dhami (2016, p. 2) opens his advanced textbook on the *Foundations of Behavioral Economic Analysis* with the following claim:

> Behavioral economics is an enhancement of neoclassical economics to take account of more empirically supported evidence on human behavior, and not its antithesis. Second, there is no paradigmatic battle between behavioral economics and neoclassical economics. As in every science, we progress by taking account of evidence that suggests a refinement and improvement of existing models.

These and many other case studies deserve a separate study, that we have to leave to a future occasion. In this connection, we acknowledge that the present chapter is only a first step toward a fully defensible account of the method of idealization and de-idealization in economics, let alone in other scientific disciplines. To this purpose, one should better develop and defend the definition of de-idealized models presented in Section 2.5; study the logical properties of the relation "M_2 is a de-idealization of M_1," where M_1 and M_2 represent scientific models in general; and explore its applicability to further case-studies in different domains besides IO theory. Again, these are all topics whose discussion is left for the future.

Finally, in this chapter we remained entirely silent on some crucial philosophical issues in economic methodology, to which our argument is clearly relevant. One has been much discussed in recent literature, and concerns the virtues of pluralism in economic theorizing and practice. In the aftermath of Rodrik's *Economics Rules* (Rodrik 2015), many scholars have strongly defended the advantages and desirability of a greater pluralism in the methods, aims and topics of current economic research. Among them, several philosophers of economics started focusing their attention on families of models, rather than on single models, in order to discuss interesting methodological issues (Aydinonat 2018; Grüne-Yanoff and Marchionni 2018; Gräbner

and Strunk 2020; Lisciandra and Korbmacher 2021). Despite their different views on economic modeling, the shared idea is that knowledge accumulation in economics mainly proceeds "horizontally", i.e., by developing many different and sometimes mutually inconsistent models of the same target, each of which is only partial and highly context-sensitive. Here, we do not need to take stock of this discussion. What is worth noting, however, is that whereas the horizontal dimension of theoretical development in economics is clearly important, the "vertical" one should not be forgotten. As we argued in this chapter, a crucial way in which economic modeling proceeds is the development of less idealized models, which relax some of the idealized assumptions of their predecessors. This vertical growth of (families of) economic models can take various forms (recall the discussion of Reiss's Hegelian metaphor in the last section), but it is clearly as fundamental as the horizontal one. For this reason, philosophers of economics attentive to scientific practice should study both types of theoretical development, and the notion of de-idealization discussed in this chapter seems a fruitful tool to do this.

A second, interesting question concerns the issue of the "credibility" of economic models. Economist Robert Sugden (2000, 2009) introduced the notion of credibility in the philosophical literature, arguing that economic models depict credible but counterfactual worlds which however are sufficiently similar to some target, real-world situation. As Sugden (2009, p. 4) himself acknowledges, the notion of "credible model" is inherently vague and it has prompted a great deal of debate among philosophers of economics (cf. Grüne-Yanoff 2009). An interesting open issue for future work is to explore whether the notion of de-idealization discussed here can contribute to the literature on the credibility of economic models. Our conjecture, that we cannot defend here in detail, is that the de-idealization strategies discussed in this chapter may at least partly explain the credibility of economic models. Indeed, if economists think that models can be increasingly made more adequate and realistic, this may in turn increase their confidence in at least some of them as reasonable depictions of the target domain. However, corroborating such a claim — i.e., that de-idealizing a model may lead, or even generally leads, to an increase in its credibility — would require notions of both de-idealization and credibility that are better and more rigorously defined than those available at present.

The credibility issue leads us to a third, and final, problem: that of realism and anti-realism. This is of course crucial also because, as recalled in Section 2.2, philosophers leaning toward realism have often invoked de-idealization as an important component of a realist view of scientific progress. The basic idea is that highly idealized models and theories can be improved via de-idealization, i.e., by developing less idealized models and theories which relax some of the idealized

assumptions of the former and are, in this sense, more realistic. This suggests that scientific progress is possible via such process of de-idealization, and that at least some of the historical changes in a scientific discipline can be interpreted as improvements toward models and theories which are less idealized and "closer to the truth" than their predecessors. This "verisimilitudinarian" view of progress, as based on the notion of verisimilitude or truthlikeness of theories and models (Popper 1963), has been long studied by philosophers (e.g. Cevolani and Tambolo 2013) and also applied to economics (Niiniluoto 2002, 2012). Further work is needed, however, to assess in detail the nature and role of idealized and de-idealized models in connection with the issues of progress and scientific realism, and to assess whether de-idealization strategies can contribute to the case for realism in economic methodology.

ADDENDUM. EQUILIBRIUM PRICE IN THE VARIAN MODEL

In this addendum, we show how to formally derive the equilibrium of the Varian model. Our proof is closely related to the original article by Varian (1980) and the later presentation of the model made in Varian (1992, pp. 292–294). The main difference is that we fix the number of competing firms exogenously and assume they have identical and constant marginal costs. Varian determines the number of competing firms according to a zero profit condition assuming that firms have zero marginal costs but identical fixed costs (it is precisely these 'extreme' hypotheses of this kind that are the subject of de-idealization).

Let $F(p)$ be the cumulative distribution function of the equilibrium strategy, that is the probability that a chosen price is lower or equal to p. Suppose that the firm 1 chooses p. If firm 2 chooses a price higher than p (an event with probability $1 - F(p)$), then p is the lowest price and the firm 1 attracts all informed consumers S. Instead, if firm 2 chooses a price lower than p (an event with probability $F(p)$), then firm 1 fails to have the lowest price and lose all informed consumers S. In the former case, the firm 1 gets a revenue of $(p - c)(S + U/2)$ (all informed and its part of uninformed); in the latter case, it gets a revenue of $(p - c)U/2$ (only its part of uninformed). In either case it pays a marginal cost c. Hence, the expected profits for the firm 1 are the following (the problem is symmetric for the firm 2):

$$\pi = \int_0^\infty \{(p - c)(1 - F(p))(S + U/2) + (p - c)F(p)U/2\}f(p)dp \quad (23)$$

Note that every price charged in the equilibrium must yield the same expected profits, otherwise it would be profitable for the firm to increase the frequency with which it charged the more profitable prices relative to the less profitable. This means that it must hold:

$$\pi = (p - c)(1 - F(p))(S + U/2) + (p - c)F(p)U/2 \tag{24}$$

or, rearranging,

$$F(p) = 1 + \frac{U/2}{S} - \frac{\pi}{2S(p - c)} \tag{25}$$

To determine π let us note that the probability that a firm charges a price less than or equal to the reservation price v is 1, so $F(v) = 1$. Solving this equation gives us $\pi = (v - c)U$ and substituting into (25) we get:

$$F(p) = 1 - \left(\frac{(v - p)(U/2)}{(p - c)S} \right) \tag{26}$$

The lower support of the distribution must satisfy $F(p_0) = 0$, so that:

$$p_0 = c + (v - c)\left(\frac{U/2}{U/2 + S} \right) \tag{27}$$

Intuitively, pricing below p_0 is not beneficial because it does not attract additional consumers and yields lower profits. Eventually, no price above the reservation price v will be charge because there is zero demand at any such price. Thus, v is the upper support of $F(p)$.

3

MODELS ON TRIAL: ANTITRUST EXPERTS FACE DAUBERT CHALLENGES

If we are to use effectively these abstract models and this descriptive material, we must have a comparable exploration of the criteria for determining what abstract model it is best to use for particular kinds of problems, what entities in the abstract model are to be identified with what observable entities, and what features of the problem or of the circumstances have the greatest effect on the accuracy of the predictions yielded by a particular model or theory.

— Milton Friedman[1]

3.1 INTRODUCTION

Imagine being an economic expert hired and generously remunerated by a party in an antitrust lawsuit. Your expertise has been called upon to predict the potential price increase following a proposed merger, a vital figure for proving the case in favor or against the merger. To do so, you must choose an appropriate model to represent the functioning of competition in the relevant market, build an empirical model using real-world data, and then employ it to predict the price increase after the merger. Eventually, you submit your expert report which details your methodology and results. Even before you can testify before the jury and be subjected to cross-examination you face a stumbling block: the opposing party files a motion to ask the court to assess the admissibility of your testimony, with the goal of excluding it at a preliminary stage. Following the motion to exclude, the judge must determine whether your model-based argument is reliable enough to produce results that constitute admissible testimony. If successful, the challenge may strike a decisive blow to your party's chances and to your reputation as a competent expert — an outcome you certainly want to avoid.

Such a scenario is commonplace in US federal courts, especially in those areas of law that rely on technical or scientific evidence. Models, or particularly the employment of model-based arguments by expert witnesses, are often on trial on charges of not being science. At least three actors interact in the trial, each with different goals. The expert witness tries to persuade the judge that her model-based arguments are reliable and should be admitted at trial. The opposing party seeks to prove to the judge that the expert testimony is unre-

1 "The methodology of positive economics", Friedman (1953, p. 42).

liable and should be excluded. Finally, the judge makes the decision with the aim of keeping junk science away from the jury.

Antitrust litigation provides an interesting case study because of its heavy reliance on the judicial evaluation of complex economic and econometric analysis (Cao 2022; Lopatka 2016; Shapiro 2021a).[2] This is hardly surprising as antitrust statutes are grounded in economic concepts such as monopolization, restraint of trade, and lessening of competition. For this reason, when it comes to finding whether a firm has engaged in anticompetitive behavior and violated antitrust law, economists are often called upon as expert witnesses by the parties in litigation (Mandel 1999; Posner 1999).[3]

This chapter explores how and why economic expert witnesses use models in antitrust litigation and how their model-based arguments are challenged under the so-called *Daubert* standard. To study how model-based arguments work in antitrust cases, I propose a new framework which portrays economic experts testifying in antitrust cases as making two applicability claims. The first are weak applicability claims where evidence is advanced to show that some critical assumptions of a theoretical model are shared by the intended target. The second are strong applicability claims where an empirical model (as opposed to a merely theoretical model) is designed and different sources of evidence are used to show that it describes (or predicts) some quantitative features of the target. Having clarified what kinds of evidence support the applicability claims, I explain how *Daubert* challenges work when models are at stake. Finally, I examine the decision-making process of the judge in assessing model-based arguments under *Daubert* in light of my framework and advance some proposals for more efficient *Daubert* hearings.

This chapter serves two main purposes. The first is to study the problem of model applicability in a context that has been neglected by philosophers of economics, namely the use of models by economic experts in court cases. In this regard, my work adds to the broad literature concerning how economic models are (and ought to be) employed for explanatory, predictive, or policy purposes (e.g. Aydinonat 2018; Jhun 2021; Mäki 2018; Rodrik 2015). Second, my chapter adds to recent work that has questioned the cognitive ability of judges in understanding expert testimonies in antitrust lawsuits, especially when they are confronted with complex and quantitative evidence that results from model-based arguments.[4] In contrast to the above literature, which focuses mainly on the judge's decision-making process,

2 In addition to antitrust law, other legal areas where economic reasoning finds an eminent place are securities regulation, employment discrimination, and patent infringement (see Chassonnery-Zaïgouche 2020; Maas and Svorenčík 2017.)

3 The name originates from *Daubert v. Merrell Dow Pharm. Inc.*, 509 US 579 (1993).

4 See, among others: Baye and Wright (2011), Coate and Fischer (2012), Giocoli (2015, 2020), Hovenkamp (2005), Salop (2021), Shapiro (2021b), Werden, Froeb, and Scheffman (2004), and Werden (2008).

here the emphasis rests on the model-based arguments proposed by expert witnesses.

I proceed as follows. Sections 3.2 and 3.3 survey the role of economic expert witnesses in US antitrust litigation and how the *Daubert* admissibility standard works when economic models are at stake. Section 3.4 develops a framework of model applicability in the context of antitrust litigation, which is employed in Section 3.5 to examine a recent US antitrust case where economic modeling is at the heart of a *Daubert* challenge. Section 3.6 examines further evidence from antitrust court cases and makes two proposals for more efficient *Daubert* hearings against model-based arguments

3.2 ECONOMIC EXPERT WITNESS UNDER THE DAUBERT STANDARD

The law governing the admissibility of expert witnesses in federal courts, set forth in the Rule 702 of the Federal Rules of Evidence, reflects the standard articulated by the Supreme Court in *Daubert v. Merrell Dow Pharm. Inc.* (1993) and then refined in a series of rulings – known as the *Daubert* trilogy – taken between 1993 and 1999. In the legal jargon, the *Daubert* standard and the ensuing Rule 702 have assigned courts the role of gatekeepers for the admission of scientific experts – a role they must perform by ensuring that 'any and all scientific testimony or evidence admitted is not only relevant, but reliable' (*Daubert*, p. 589). More specifically, Rule 702 states that trial judges serve a gatekeeping role to determine whether

(1) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;

(2) the testimony is based on sufficient facts or data;

(3) the testimony is the product of reliable principles and methods; and

(4) the expert has reliably applied the principles and methods to the facts of the case.[5]

The introduction of the *Daubert* standard represents a major turning point in the interaction between the judge and the expert witness because it asks the court to determine whether the expert testimony can be admitted on grounds that it is the result of scientific methodology. An important legacy of *Daubert* has been the emergence and spread of the so-called *Daubert* challenge, a special motion made to

5 Rule 702 – Testimony by Expert Witness. Retrieved from: https://www.law.cornell.edu/rules/fre/rule_702.

the judge before or during litigation in order to prevent the introduction of a specific expert testimony to the jury. *Daubert* challenges enriched the tools in a litigator's belt because a successful exclusion of the opposing party's expert testimony can potentially change the outcome of the case. Although the *Daubert* standard aimed to be "a vital weapon in the mounting war against 'Junk science' in the courts'" (Gavil 2001, p. 6), many concerns have been expressed from both law scholars and philosophers of science (e.g. Haack 2005, 2015; Martini 2015). Here I will take the *Daubert* standard at face value and focus on how economic expert witnesses deal with it when using economic models in the courtroom.[6]

The empirical evidence on the effect of the adoption of the *Daubert* standard on economic expert witnesses is rather consistent. Three results, which are robust across different empirical studies, are particularly telling (Giocoli 2020; Langenfeld and Alexander 2011)[7] First, *Daubert* challenges are successful in excluding, totally or partially, the testimonies of economic experts in around 40% of cases. Second, economists receive a relatively higher percentage of *Daubert* challenges in antitrust cases compared to other legal areas where they testify as expert witnesses. Third, plaintiffs' economic experts appear to be excluded more easily than those of defendants.[8] Although a comparative analysis on the number of challenges and exclusion rate of economists compared to other scientific disciplines is still missing, these results suggest that complying with the *Daubert* standard is a critical task facing antitrust economists in US courts.[9]

3.3 MODELING THE BUT-FOR WORLD IN ANTITRUST LITIGATION

In many antitrust cases economic experts need to model a counterfactual scenario of what would have happened in the absence of the anticompetitive conduct. Antitrust practitioners usually refer to this counterfactual scenario as the but-for world to indicate the world that one should expect to exist but for the alleged anticompetitive behavior. Ideally, the comparison between the actual world and the but-for world demonstrates the effect of the antitrust violation. Economists mostly use theoretical models from industrial organization to describe such counterfactual worlds. A notable example of but-for modeling is merger simulation, where game-theoretic models of oligopolis-

6 A critical discussion of the *Daubert* standard can be found in chapter 4.

7 These findings should however be taken with a certain degree of caution for it is extremely hard to build a dataset containing all and only *Daubert*-related challenges (see Langenfeld and Alexander 2011, Appendix.)

8 Appendix A below offers an update on the existing empirical studies on *Daubert* challenges to economic expert witnesses.

9 The *Daubert* Tracker website reports that economics ranks fourth among the most challenged discipline under *Daubert*/Rule 702 after medicine, engineering, and psychology. See www.dauberttracker.com (accessed May 15, 2023).

tic markets are employed to predict what would happen to prices in the relevant market if the merger were allowed.[10]

When economic models are at stake, federal judges and practitioners have interpreted the *Daubert* standard as a "fit requirement" and the literature on economic experts in antitrust has embraced this terminology. The fit requirement is a normative criterion according to which economic models employed by expert witnesses should fit the particular facts of the case in order to satisfy the *Daubert* standard. Gregory Werden, a leading antitrust scholar, explains the fit requirement as follows:

> Economic modeling of the sort that may be relied on by an expert in an antitrust case entails choices made by the expert, and most choices should be justified on the basis that they are consistent with the facts of the case. Evaluating the fit of an economic model draws on the full array of qualitative and quantitative evidence developed in the case. Courts have excluded economic testimony in antitrust cases premised on models not sufficiently grounded in the facts. [...] An empirical model used to make predictions or disentangle effects also must be tailored to the facts of the case. Werden (2008, p. 811)

Indeed, the failure of economic models to fit the facts of the case has been frequently mentioned by judges as a compelling reason to exclude expert testimonies under the *Daubert* standard.[11] As a result, debate among antitrust scholars centers upon whether the *Daubert* standard – interpreted as a fit requirement – is an appropriate normative requirement for economics (Hovenkamp 2005; Lopatka 2016; Werden, Froeb, and Scheffman 2004; Werden 2008 and what families of models are more likely to survive a *Daubert* challenge (Coate and Fischer 2012; Giocoli 2015, 2020).

Unfortunately, the fit requirement is couched in rather fuzzy terms and leaves many questions unanswered. First, it is not entirely clear what "model fitness" amounts to. Does model fitness refer to a sort of similarity between the theoretical model and the target or to the successful predictive performance of an empirical model based on the theoretical one? An economic model may fit the target in the former

10 Merger simulation has been introduced as a methodology for antitrust enforcement in the mid-1990s and it is now considered a standard tool by antitrust authorities. For a survey on merger simulation, see Budzinski and Ruhmer (2010).

11 Stanford economist Robert Hall saw his testimony rejected in *Concord Boat Corp v. Brunswick* (207 F.3d 1039, 8th Cir., 2000) because the Cournot duopoly model proposed to describe the relevant market did not fit economic realities (see Werden, Froeb, and Scheffman 2004, p. 89). Three years later, a federal judge excluded the plaintiff's expert following a *Daubert* challenge on similar grounds in *Heary Bros. Lightning Prot. Co., Inc. v. Lightning Prot. Inst.*, 287 F. Supp. 2d 1038 (D. Ariz. 2003). See *infra*, Section 3.6. On Hall's exclusion, see below, chapter 4 of this dissertation.

sense but not in the latter. Does such a model satisfy the fit requirement? At the present stage, we are left without an answer. Second, the fit requirement does not specify the relations between model fitness and the structure of *Daubert* challenges against model-based arguments, that is, it does not explain how *Daubert* challenges work when model fitness is at stake. For instance, what kinds of evidence are put forward by experts in their model-based arguments and are typically challenged under *Daubert*? Again, we are left without an answer. Given the importance of model-based arguments in antitrust cases and the widespread reliance on *Daubert* challenges, an alternative framework is much needed.

3.4 MODEL APPLICABILITY IN ANTITRUST CASES: A NEW FRAMEWORK

Model fitness usually refers to how well a statistical model explains the available data. Such a narrow concept does not help us understand the complex entanglement between theoretical and empirical models, evidence, and *Daubert* challenges that shape antitrust court cases. Following recent work in the philosophy of economics, I spell out the relationship between economic models and reality in antitrust court cases in terms of *applicability*.

The issue of what it means to apply a model to a real-world target is a traditional one in the philosophy of economics (e.g. Gibbard and Varian 1978; Hausman 1992). In a recent twist in this debate, Rodrik's *Economics Rules* (2015) has emphasized several issues involved in model applicability from the point of view of a practitioner. His starting point is the criticism that economic models are riddled with false assumptions (for this reason they are called 'unrealistic' or 'idealized': Hindriks 2006; Mäki 2020; Peruzzi and Cevolani 2022) and, therefore, that such models cannot be properly applied for explanatory or policy tasks.[12] Rodrik proposes that the applicability of a model depends not only on model-target resemblance but rather on the purpose of the modeler, the context of the application, and the audience being addressed. This is because purpose, context, and audience, determine which assumptions are critical in the given model application and – as he contends – 'what matters to the empirical relevance of a model is the realism of its critical assumptions' (Rodrik 2015, p. 55)

Rodrik's own example of the perfectly competitive market model clarifies this point (pp. 21–22). Suppose an economist wants to apply the perfectly competitive model in order to advise the government about the effect of imposing price controls on the cigarette industry. In a perfectly competitive market model, a price control would lead to

12 The classical reference on the topic of the realism of assumptions in economics is Friedman (1953) and the vast debate following it (see Mäki 2009).

a decrease in the supply of cigarettes. Given the purpose, context, and audience of the modeler, the assumption that firms have no market power is critical because a modification of that assumption would produce an extremely different result (for instance, a price cap in a monopolized market would induce the firm to increase its output). At the same time, other assumptions concerning the exact number of firms and the agents' degree of rationality would not overly affect the perfectly competitive model's results and, therefore, are not critical for the given application. Given that 'firms have no market power' is a critical assumption for this specific application, economists should attempt to check whether this assumption is approximately satisfied in the cigarette industry.

The above discussion on model applicability is strictly related to a further topic that has been widely debated by philosophers of economics, namely model selection. Following Aydinonat 2018, model selection can be understood as a procedure to select from a model library a single or a set of models for a specific target for a particular purpose. How are the notions of model selection and critical assumption related? To put it simply, candidate (sets of) models are distinguished and selected on the basis of their critical assumptions (Grüne-Yanoff and Marchionni 2018; Rodrik 2015. Given the purpose of predicting the effect of price controls on the cigarette industry, the perfectly competitive model can be selected as an applicable one if, among other things, the critical assumption about firms' market power is realistic enough for the given situation. In other words, identifying critical assumptions and judging 'the extent to which they approximate the particular situation to which a model is supposed to apply' are important steps in the model selection procedure (Grüne-Yanoff and Marchionni 2018, p. 3).

3.4.1 *Weak applicability claims*

Model selection based on critical assumptions is an important part in expert testimony because antitrust practitioners must decide which theoretical model of competition best characterizes a given market. I argue that the first activity performed by the economic expert aims at establishing a *weak applicability claim*:

- A theoretical model M is weakly applicable to a target T because there is evidence $E_1, \ldots, E_n$ that some critical assumptions of M are approximately shared by T.

In the context of antitrust cases, the evidence $E_1, \ldots, E_n$ consists mainly of qualitative and informal evidence, with a high degree of subjective judgment by the economic expert. The goal of the expert is to convince the judge that the selected model shares important fea-

tures with the intended target, usually a very specific market limited in space and time.

Suppose, for example, that you are the economic expert hired by the plaintiff to compute the but-for prices in the North Carolina fruit market in the absence of firm A's anticompetitive conduct. Let us consider the but-for North Carolina fruit market as the target and assume, for simplicity, that firm A faces a single competing firm B. As a first step you must select a theoretical model by providing evidence that some critical assumptions of the model are approximately shared by the target given your purpose. For example, you might observe that products sold by the two firms are not close substitutes and that firms have negligible capacity constraints. These two features of the market lead you to discard models that assume product homogeneity and quantity competition. In the end, you might decide to use a D-Bertrand model with differentiated goods (henceforth, D-Bertrand model) because evidence exists that some critical assumptions (product differentiation and price competition) are realistic enough given the target.[13]

Once you have chosen a theoretical model of competition, your job as expert witness is not over because you still need to employ the model to describe or predict quantitative features of the target – in this case the but-for prices in the North Carolina fruit market. To do that, you must decide how to give particular values to the model's parameters using real-world data and then provide evidence that your estimates are reliable. To account for this particular use of economic models in antitrust litigation and its relation with *Daubert* challenges, I will introduce the notion of *strong applicability claim*.

3.4.2 *Empirical models and strong applicability claims*

While weak applicability can be defined solely in terms of theoretical models, strong applicability is to be defined using the notion of *empirical models*. To provide a thorough explication of the notion of empirical models is beyond the scope of this chapter. For my purposes, the following working definition will suffice:

- An empirical model M' is the result of a series of activities that provide inputs to, modify, and augment the theoretical model M in order to obtain numerical results from it in connection to the expert's goal.[14]

Let us consider again the North Carolina fruit market modeled as a D-Bertrand market. The D-Bertrand model's equilibrium result is a

13 For an in-depth analysis on the main characteristics of the D-Bertrand model, see chapter 2, Section 2.4.

14 Economists usually refer to the procedure of building empirical models of such kind as structural modelling approach (e.g. Reiss and Wolak 2007).

set of first-order conditions that arises from firms' profit maximization. First-order conditions determine a relation between three variables: market prices, marginal costs, and demand elasticities. Since the expert's goal is to solve for equilibrium prices of the target market, she needs to find numerical values for marginal costs and elasticities of demand. When numerical values are substituted for the model's parameters, the economic expert is no longer dealing with the theoretical D-Bertrand model but rather with an empirical version of the D-Bertrand model.[15]

Using the notion of empirical model, I am ready to define a *strong applicability claim*:

- An empirical model M', based on the theoretical model M, is strongly applicable to the target T because there is evidence $E'_1, \ldots, E'_m$ that M' describes or predicts some quantitative features of the target to some degree.

A strong applicability claim thus involves two intertwined steps, namely the design of an empirical model and the search for evidence that the empirical model describes or predicts quantitative features of the target. Broadly speaking, antitrust economic experts build empirical models by specifying a functional form of demand, estimating the model's parameters using real-world market data, and adjusting theoretical assumptions to get closer to the intended target. However, no ready-to-use recipe exists given that every case has its own specificity that depends crucially on the context, purpose and audience.[16]

As to what constitutes evidence $E'_1, \ldots, E'_m$, let us consider the case of a merger simulation. After having predicted the post-merger prices using an empirical model, the expert witness cannot directly check whether her price estimates are borne out by real-world data simply because the merger has not yet happened and may never happen if not cleared. This situation is endemic because, as we have seen, expert witnesses are often called to model counterfactual scenarios for which no direct evidence is available.

I classify the evidence $E'_1, \ldots, E'_m$ for strong applicability claims in three groups: historical evidence, indirect evidence, and sensitivity analysis.[17] Using historical evidence means that the expert, whenever possible, should compare the numerical results of the empirical

15 In the field of industrial economics, the distinction between game-theoretic models that represent market competition in the abstract and their empirical counterparts, which result from parametrizing such models using real-world data, is relatively clear. However, when we consider economics in a broader context, delineating a precise line between theoretical and empirical models becomes a challenging task and it remains an open question in the philosophy of economics.

16 Section 3.5 *infra* will discuss a concrete example of an empirical model designed by the economic expert witness in an antitrust dispute.

17 The three kinds of evidence are not mutually exclusive. For a concrete example of an expert witness using both indirect evidence and sensitivity analysis, see Section 3.5 below.

model with historical data available for the target. For example, it is customary to require that empirical models in merger simulations account for the average level of prices over a year in the industry under review (Werden, Froeb, and Scheffman 2004, p. 90). Indirect evidence, by contrast, is suitable whenever the target of the empirical model is a counterfactual but-for world for which no historical data are available. In such cases economic experts may use indirect evidence from other sources to convince the judge of the reliability of their empirical model's result. Suppose the economic expert's empirical model generates but-for market prices for apples around 10 dollars and that additional evidence exists that the average price of apples in similar markets varies from 8 to 13 dollars. This additional evidence may be brought by the expert to strengthen her strong applicability claim (see Section 5 for an example). Eventually, economic experts may use sensitivity analysis to make clear how alternative choices of functional forms and parameters affect the empirical model's result (Woodward 2006). A typical example consists in estimating elasticities of demand from the same body of data but using a variety of different assumptions about the functional form for the demand system and seeing how price predictions respond to the change of values for the elasticities of demand.

3.4.3 *Daubert challenges to model-based arguments*

So far I have depicted economic experts testifying in antitrust cases as making two applicability claims. The first is a weak applicability claim where evidence is advanced to show that some critical assumptions of a theoretical model are shared by the intended target. The second is a strong applicability claim where an empirical model is designed and different sources of evidence are used to show that it describes or predicts some quantitative features of the target.

In general, weak applicability serves as a precondition for strong applicability. This means that if an economic expert fails to prove a weak applicability claim, then she also fails to prove a strong applicability claim. This is because, as we have seen with Rodrik's examples, empirical models typically cannot yield accurate quantitative descriptions or predictions unless the underlying theoretical model shares critical assumptions with the target. For example, an empirical version of the D-Bertrand model can predict accurately certain aspects of the target market only if the theoretical D-Bertrand model captures the critical features of the competitive process in that market. Also, evidence from *Daubert* hearings demonstrates that judges consider weak applicability as a precondition for strong applicability. In other words, if judges determine that an economic expert's model

lacks weak applicability, they do not proceed to evaluate its strong applicability (see Section 3.6 for further details).[18]

Let us now employ this framework to explain what most *Daubert* challenges against model- based arguments are about, viz. the strength of the evidence that the expert witness puts forward to justify the two applicability claims. Expert witnesses are usually highly qualified and employ models that are generally accepted by the scientific community. However, this does not prevent them from receiving *Daubert* challenges that cast doubt on the evidence $E_1, \ldots, E_n$ and $E'_1, \ldots, E'_m$ advanced for the two applicability claims. This can be done either by directly criticizing the evidence offered by the expert witness or by providing alternative evidence. Using our previous example, a *Daubert* challenge might be based on evidence showing that firms have significant capacity constraint, thus implying that the D-Bertrand model is not weakly applicable to the target. The same challenge could also offer a detailed critique of how marginal costs should have been inferred from accounting data or why demand elasticities are poorly estimated, thus implying that the empirical D-Bertrand model is not strongly applicable to the target.

More formally, a *Daubert* challenge against model-based arguments consists of two main sub-challenges to the opposing party's expert testimony:

- Weak applicability sub-challenge: the evidence $E_1, \ldots, E_n$ is insufficient to justify the weak applicability claim by the expert witness.
- Strong applicability sub-challenge: the evidence $E'_1, \ldots, E'_m$ is insufficient to justify the strong applicability claim by the expert witness. is insufficient to justify the strong applicability claim by the expert witness.

The following section will show that my framework is descriptively adequate by reviewing a 2015 US antitrust case where economic modeling was at the heart of a *Daubert* challenge.

3.5 CASE STUDY: CASTRO V. SANOFI PASTEUR INC. (2015)

A 2015 antitrust case, *Castro v. Sanofi Pasteur Inc.*, centered around economic expert testimony and its admissibility under the *Daubert* standard. For this case, direct documentation exists of the report of

18 It is important to note that the weak applicability of a theoretical model does not guarantee the strong applicability of an empirical model derived from it. So, that the critical assumptions of the D-Bertrand model are shared by the North Carolina fruit market does not imply that the predictions of the empirical model will be accurate (for example, if demand elasticities are poorly estimated). Therefore, once federal judges are satisfied with the weak applicability claim, they proceed to examine the strong applicability claim (see below, section 3.6).

the plaintiff's expert, the *Daubert* challenge raised by the defendant, and the district court's opinion.[19]

Castro concerns the antitrust action raised by a number of physicians, led by Dr. Adriana Castro, against the bundling of pediatric vaccines by the French multinational pharmaceutical company Sanofi. Plaintiff's expert, renowned antitrust scholar and Harvard professor Einer Elhauge, was testifying at the class certification stage, but the defendant raised a motion to exclude his opinion under *Daubert* / Rule 702. The District Court of New Jersey conducted a *Daubert* hearing to evaluate whether or not to accept the challenge. Eventually, the court was convinced that the expert testimony did satisfy the *Daubert* standard and, therefore, could not be excluded from the trial. I will examine this case on the basis of the framework developed in the previous section.

3.5.1 *Plaintiff's expert report*

Sanofi had for many years monopoly power in the market for the conjugate quadrivalent meningococcal vaccine (MCV4), which is usually used to inoculate children against meningitis bacterium. By mid-2009, Sanofi became aware that its own MCV4 vaccine, Menactra, would have soon faced the competition of a new vaccine, Menveo, produced by market entrant Novartis. Sanofi responded to the entry threat by bundling Menactra with its other pediatric vaccines and substantially increasing its prices (*Castro*, p. 826). Only customers who purchased the whole bundle of vaccines were able to receive a 'loyalty discount' that dropped prices back to the level preceding Menveo's entry into the market. All other customers paid higher prices on all vaccines in the bundle. The plaintiffs were three pediatric physicians who sought to represent a broader class of individuals and firms that were directly damaged by Sanofi's behavior which led to higher prices for vaccines.

Following this antitrust allegation, Elhauge's expert report had to provide evidence about what the prices paid by consumers would have been absent the bundle. In other words, he was required to model a but-for world that should represent the market for MCV4 vaccines in the absence of Sanofi's alleged anticompetitive conduct. He described his task as follows:

> One can estimate how much the Bundle has increased MCV4 prices above but-for levels (i.e. levels that would have prevailed absent the Bundle) by comparing actual MCV4 prices with the Bundle to what an economic model

19 *Castro v. Sanofi Pasteur Inc.*, 134 F. Supp. 3d 820 (D.N.J. 2015). I accessed the court's opinion, legal briefs, motions, and other unpublished material through the legal database *LexisNexis*.

> based on actual market data indicates the firms would have charged without the Bundle. (Exhibit 51B, p. 147)[20]

A substantial portion of his expert's report is devoted to the task of determining the but-for prices and comparing them with actual prices, to show that consumers were paying inflated prices as a result of the bundle. The report is divided into different parts:

> Part A below explains that the most appropriate model to use in this market is the "differentiated Bertrand competition" model. [...] Part B describes the quantitative market data used as inputs to the model. Part C calculates but-for profit-maximizing Menactra and Menveo prices by inputting the quantitative market data into the differentiated Bertrand model. (Exhibit 51B, pp. 147–148)

Part A roughly reflects the activity of making a weak applicability claim, while parts B and C correspond to the design and justification of an empirical model, that is, a strong applicability claim. Let us begin with the former.

3.5.1.1 *Weak applicability claim*

Elhauge chooses to model the but-for world using a Bertrand model with differentiated goods, a standard model of industrial organization theory widely used in the antitrust context.[21] Roughly speaking, the D-Bertrand model assumes that firms produce differentiated products and compete by setting prices in order to maximize their profits. For this reason, the plaintiff's expert strives to provide evidence that the two MCV4 vaccines – Menactra and Menveo – differ in a significant manner for consumers and that the two firms compete on prices. For example, Elhauge argues that the two vaccines have different chemical properties and are approved for different age groups, thereby prompting some medical providers to prefer one to the other (Exhibit 51B, p. 149). At the same time, internal documents show that neither Sanofi nor Novartis have significant capacity constraint in the MCV4 market. If no capacity constraint is in force, then price competition seems to be more appropriate than quantity competition: firms set prices and are able to sell as much output as demanded by buyers at offered prices (Exhibit 51B, pp. 155–156).

In the terminology of Section 3.4, the economic expert is providing evidence $E_1, \ldots, E_n$ to show that critical assumptions of the theoretical D-Bertrand model (product differentiation and price competition) are shared by the intended target, that is, the US market for MCV4

20 Castro, Exhibit 51B. Retrieved from: https://advance.lexis.com/api/document?collection=briefs-pleadings-motions&id=urn:contentItem:5KS1-D4J1-JNCK-21KY-00000-00&context=1516831.

21 For a textbook treatment of the model, see Tirole (1988).

vaccines in the period 2010–2015. Having provided such evidence, Elhauge (Exhibit 51B) concluded that 'market characteristics and market data both indicate that, without the Bundle, Sanofi and Novartis would engage in "differentiated Bertrand competition" in the private segment of the MCV4 market' (p. 148).

3.5.1.2 *Strong applicability claim*

Given the purpose of obtaining numerical values for market prices of the two vaccines (Menactra and Menveo) produced by Sanofi and Novartis respectively, the economic expert needs to provide numerical inputs for both firms' marginal costs and consumers' demand elasticities. In Elhauge's words:

> The next step in estimating the but-for MCV4 prices is to calibrate the differentiated Bertrand competition model to the MCV4 market by obtaining quantitative estimates – based on the actual market data, internal company documents, and other information – of the key factors that determine profit-maximizing MCV4 prices. (Exhibit 51B, p. 168)

Researchers and practitioners routinely use a variety of approaches to measure marginal costs. One of these approaches is based on accounting data, that is, to leverage internal firm documents and determine which costs are most likely to stand for the theoretical notion of marginal costs. Elhauge's expert report follows this approach.

As often happens in the context of court litigation, some experts' choices are dictated by strategic considerations related to the goal of persuading the judge of the merits of their analysis. In this case, Elhauge decided to consider as marginal all cost categories that Sanofi includes as 'direct cost of good sold' in its balance sheet (Exhibit 51B, p. 171). The choice tends to inflate Sanofi's marginal costs for its vaccine Menactra and generate, ceteris paribus, higher but-for prices. This conservative approach is strategic because it suggests to the judge that the plaintiff's expert is not trying to artificially inflate the difference between actual prices and but-for prices.

Besides marginal costs, Elhauge had to estimate own- and cross-price elasticities for Menactra and Menveo. To explain how antitrust economists model demand systems would require an extensive detour (Björnerstedt and Verboven 2016; Budzinski and Ruhmer 2010; Werden, Froeb, and Scheffman 2004). Suffices here to say that in *Castro* Elhauge estimated demand elasticities using a discrete choice logit model and feeding it with actual data from a portion of the MCV4 market unaffected by the challenged conduct.[22]

22 To be precise, he identified such portion of the market in the Federal Supply Schedule (FSS), viz. long-term contracts with companies that provide access to commer-

The empirical model is solved and but-for prices are computed. Elhauge found that but-for prices for both competing vaccines Menactra and Menveo would have been substantially lower than actual prices. For example, Menactra actual price in 2013 was $99.36 while its but-for price derived from Elhauge's empirical model would have been $61.25. The 38% difference between the actual and the but-for price was the overcharge that could be imputed to the alleged anti-competitive bundling (Exhibit 51B, p. 204).

To show that his empirical model successfully predicted the but-for prices Elhauge exploited both indirect evidence and sensitivity analysis. For example, a section called 'Other evidence confirms plausibility of these but-for prices' argues that but-for prices and the relative profit margin were both within normal range for the vaccine industry (Exhibit 51B, pp. 205–206). In addition to indirect evidence, Elhauge performed sensitivity analysis concerning marginal costs – that is, he studied how different numerical values for marginal costs would affect the values of but-for prices. His conclusion was that alternative choices would give rise to lower but-for prices (and thus even higher overcharges), thus reinforcing the idea that his but-for prices were a conservative result (p. 185).

3.5.2 *The Daubert challenge*

Section 3.4 argued that many *Daubert* challenges aim to cast doubt on the evidence $E_1, \ldots, E_n$ and $E'_1, \ldots, E'_m$ that the expert uses to put forward the two model applicability claims. The *Castro* case is no exception. Sanofi's attorneys raised a motion to exclude Elhauge's testimony under the *Daubert* standard:

> The question presented by Sanofi's *Daubert* motion is whether the methodology that led Elhauge to this opinion passes muster. As Sanofi explained, each of its components departs from generally accepted practices, is riddled with error, and lacks the scientific validity required by a court of law. (Reply Brief, p. 2)[23]

The first part of the challenge revolves around the fact that the plaintiff's expert 'did not ensure that the assumptions that generated the models' results were satisfied here' (Reply Brief, p. 4). This part corresponds to what I called a weak applicability sub-challenge, stating that evidence $E_1, \ldots, E_n$ is insufficient to justify the weak appli-

cial products and services at regulated prices to the government. As explained by Elhauge (Exhibit 51B), he used data on FSS purchasers because 'FSS customers were not subject to the Bundle and thus their decisions could not be distorted by the Bundle' (p. 190).

23 *Castro*, Reply Brief. Retrieved from: https://advance.lexis.com/api/document?collection=briefs-pleadings-motions&id=urn:contentItem:5YX8-9GF1-FBV7-B32M-00000-00&context=1516831.

cability claim proffered by the expert. We saw that Elhauge relied on certain evidence to show that some critical assumptions of the D-Bertrand model were shared by the target. The defendant's attorneys, however, pointed out that other assumptions of the model were highly suspicious given the intended target:

> It is the one-shot game assumption that is unrealistic. In a one-shot game, each firm has "one shot" to earn the business at stake. [...] Bertrand assumes away the most likely business strategy and the one Novartis employed in the actual world: enter at approximate parity and wait and see what happens. (Reply Brief, p. 12)

In other words, the defendant was trying to show that additional evidence existed that some critical assumptions of the D-Bertrand model were not shared by the target and that, therefore, this model was not weakly applicable to it. Defendant's attorneys seemed to suggest that alternative theoretical models where firms repeatedly interact in the product market would have been better to describe this specific vaccine industry.

The *Daubert* challenge also contains a strong applicability sub-challenge that evidence $E'_1, \ldots, E'_m$ is insufficient to justify the strong applicability claim proffered by the expert. In *Castro* this sub-challenge takes the following form:

> A simulation is not reliable "without validation," or the process of ensuring it "accurately predicts" real-world behavior. "Calibrating" the model is so important that an "inadequately calibrated" model must be "excluded under *Daubert*" because the "reliability of an economic theory is tested by comparing it to reality." (Reply Brief, p. 14)

Among the deficiencies spotted by the defendant, an important one was the fact that the plaintiff's expert allegedly used the wrong cost data to identify numerical values for marginal costs. Without this error, the defendant argued, the empirical model built by the plaintiff's expert would give rise to different but-for prices and, therefore, different overcharge (Reply Brief, pp. 18–19). For this reason, the defendant argued that the evidence proposed by plaintiff's expert Elhauge did not suffice to make a strong applicability claim and, therefore, his testimony should be excluded under the *Daubert* standard (Reply Brief, p. 20).

As already mentioned above, the district court denied the *Daubert* motion to exclude Elhauge's expert witness testimony. The court argued that even though other approaches could have been employed, including alternative theoretical models, 'the use of the differentiated Bertrand model here, however, is not error' since the plaintiff's expert 'notes that Sanofi and Novartis documents indicate there are no

production limits. And Menactra and Menveo are differentiated products' (*Castro*, p. 838). Therefore, the court was satisfied by the evidence put forward for the weak applicability of the D-Bertrand model to the but-for market for MCV4 vaccines:

> Professor Elhauge has provided defensible reasons for not using the various models Defendant proposes and all alternate models need not be ruled out. Defendant's disagreement with the model selected is not a basis for inadmissibility here. (*Castro*, p. 839, emphasis added)

The court also argued that Elhauge had built a reliable empirical model using real-world data on market demand and cost data, thereby concluding that "defendant's motion to exclude Professor Elhauge's expert reports is denied" (*Castro*, pp. 840–842).

This case is just one example of how my framework can be applied to examine model- based arguments in court cases and *Daubert* challenges against them. The next section will consider analogous cases where the use of economic models by expert witnesses has been questioned under the *Daubert* standard and advances some proposals for more efficient *Daubert* hearings.

3.6 HELPING THE JUDGE ASSESS MODEL-BASED ARGUMENTS: TWO PROPOSALS

In the previous sections economic experts and their models took center stage. Yet let us bear in mind that the federal judge has the final word in a *Daubert* hearing. As required by *Daubert*, federal judges must determine whether the methodology upon which a given expert economist relies is truly scientific and can be presented before the jury. In this context, model applicability claims seek to persuade the judge that the expert testimony is sufficiently reliable to pass muster under *Daubert*. At the same time, the party opposing the expert economist will try to show that she has not satisfied the *Daubert*'s requirements for admissibility. Putting the decision-making process of the judge back at centerstage helps us understand both the relationship between weak and strong applicability claims and the normative scope of my framework.

The evidence from antitrust cases where model-based arguments were challenged in a *Daubert* hearing reveals that federal judges enjoy some degree of discretion in their decisions. In some instances, judges have deferred the analysis of model applicability claims to cross-examination.[24] In other cases, economic experts have been excluded because of insufficient evidence for the weak applicability

24 District judge Colleen McMahon's recent comment on a *Daubert* motion provides an example. "This is a classic example", she wrote, "of what this Court calls, 'That expert's testimony hurts our case, so let's try to disqualify the expert' use of *Daubert*. A *Daubert* inquiry is designed to weed out unreliable methodologies – 'junk science'

claim, without further examination of strong applicability.[25] In line with my framework, judges review strong applicability only when they believe that weak applicability has been fulfilled. This is the case of *Castro* where the judge took the time to assess the evidence for both weak and strong applicability claims and determine whether it was sufficient or not to survive a *Daubert* motion.

A similar scenario occurred in a recent *Daubert* hearing where the evidence for a strong applicability claim was assessed by the judge because the defendants did not question weak applicability. In November 2022 consumer plaintiffs brought Google to trial for illegally monopolizing the Android app market with anticompetitive practices in the Google Play Store. An economic expert was hired to quantify the damages caused by Google's violations of federal antitrust laws.[26] The plaintiff's expert used a Rochet–Tirole model of two-sided markets to figure out the prices consumers would have paid for apps on the Play Store in a counterfactual world absent the challenged conduct. Instead of challenging the choice of the theoretical model, Google objected to the methodology used by the plaintiff's expert in estimating the pass-through rate, an essential input of his empirical model. Google argued that the accuracy of the empirical model's results was severely undermined due to the estimation error.[27] At the *Daubert* hearing, the court denied the *Daubert* challenge and postponed a more thorough analysis of strong applicability to cross-examination.[28]

–– not to be a substitute for cross-examination about the validity of an analysis that uses established methodologies. [...] Economic bargaining models have been upheld as reliable in other antitrust cases. [...] Defendants are free at trial to explore Dr. Vogt's application (or misapplication) of the model" (*In re Namenda Indirect Purchaser Antitrust Litig.*, 338 F.R.D. 527 (S.D.N.Y. 2021), p. 546).

25 In a section called 'Mr. Guth misapplied the Cournot model', judge Roslyn Silver wrote that 'if firms compete on price, the Cournot model does not apply. It is undisputed that LPS firms compete on the price, not quantity, because they compete by price bidding. [...] thus the Cournot model does not fit the economic reality' (*Heary Bros. Lightning Prot. Co. v. Lighning Prot. Inst.*, 287 F. Supp. 2D 1038 (D. Ariz. 2003), p. 1060). Likewise, in Food Lion, LLC v. Dean Foods (*In re Se. Milk Antitrust Litig.*), the district court excluded the testimony of the plaintiff's economic expert: the court claimed that the expert's Cournot model failed to consider crucial commercial realities of the relevant market (2012 U.S. Dist. LEXIS 37650, 2012-1 Trade Cas. (CCH) P77,948, 2012 WL 947106, U.S. Dist. Ct. E.D. Tenn., Greenville Div. March 20, 2012, Filed).

26 *In re Google Play Store Antitrust Litig.*, 2022 U.S. Dist. LEXIS 213670, 2022 WL 17252587 (United States District Court for the Northern District of California November 28, 2022, Filed).

27 The pass-through rate, which is determined by estimating the market demand, reveals the extent to which app developers pass on the supracompetitive costs imposed by Google to end consumers.

28 'Dr. Singer posits, without objection by Google, that the Android App Distribution Market is a two-sided market. [...] Dr. Singer said that he used for the two-sided Android App Distribution Market the Rochet–Tirole model. [...] the "pass-through" rate is a critical element of Dr. Singer's overcharge analysis, and is the main point of contention in the *Daubert* dispute. [...] Overall, Google has not demonstrated that

Daubert hearings which require the examination of strong applicability claims, such as those in *Castro* and *Google Play Store*, often involve highly technical discussions that may be challenging for judges without specialized knowledge in industrial organization and econometrics. Consequently, there is a risk that judicial decisions made in these hearings may not fully capture the complexities of the case at hand. On top of that, the fit requirement offers no clear guidance to judges as to how to assess model-based arguments because it does not specify what model fitness amounts to. Federal judges often find themselves in a dilemma, torn between assuming the role of scientific reviewers without having the necessary expertise or deferring the analysis of the expert's methodology to cross-examination, thus forgoing their role as gatekeepers. Drawing upon the framework developed thus far, I will now outline two proposals to make *Daubert* hearings more efficient.[29]

To make life easier for judges, a solution too rarely adopted is for the court to appoint its own neutral economic expert under Rule 706 of the Federal Rules of Evidence. Proposals for the increased use of court-appointed experts in complex civil litigation have been forcefully made by Posner 1999 and Sidak 2013. Court-appointed experts, they argue, are better equipped than judges to understand expert economic testimony and evaluate whether is it grounded on scientific methodology. My analysis of *Daubert* hearings where judges are presented with model-based arguments bolsters the case for wider use of court-appointed experts. We have seen that assessing the reliability of model-based arguments requires understanding what the critical assumptions of theoretical models are and whether they match the characteristics of the target market. Moreover, it may require an appreciation of how empirical models are constructed and the kinds of evidence supporting quantitative descriptions and predictions. Arguably, such an assessment exceeds the economic expertise expected of most federal judges. Therefore, when model-based arguments are involved in a *Daubert* hearing, court-appointed experts may help judges in assessing the evidence for the applicability claims and excluding testimonies that are not reliable.[30]

My second proposal is to eliminate the requirement for strong applicability at the *Daubert* hearing stage. More precisely, I recommend that judges evaluate only weak applicability claims at this stage, deferring the examination of strong applicability until cross-examination. This proposal draws upon the different levels of cognitive effort re-

unreliability or invalidity warrant exclusion of Dr. Singer's opinions' (*In re Google Play Store Antitrust Litig.*, p. 8 ff.).

29 Roughly speaking, a *Daubert* hearing is efficient when the court successfully manages three often conflicting needs: excluding unreliable scientific evidence, upholding the jury's fact-finding role, and minimizing the time and expenses incurred by all parties involved.

30 Yet, in the next chapter we will qualify this proposal.

quired to evaluate weak and strong applicability claims. Intuitively, assessing whether the predictions of empirical models are reliable enough given the array of evidence – historical, indirect, and sensitivity analysis – brought by economic experts demands more cognitive effort than assessing the similarity between theoretical model assumptions and the relevant market. If adopted, this proposal would allow judges to admit expert testimony that meets weak applicability, while still upholding the current standard of excluding testimony that fails to meet this criterion. As a result, cases such as *Castro* and *Google Play Store* would be considerably expedited, leading to substantial time and cost savings.[31]

3.7 CONCLUDING REMARKS

In this chapter I investigated the use of models by economic expert witnesses in the courtroom by providing a framework of model applicability in the context of antitrust litigation and, specifically, of *Daubert* challenges raised against a model's 'scientificity'. I employed my framework to carry out a review of *Castro* (2015), a US antitrust case where the economic expert's testimony successfully faced off a *Daubert* challenge. By looking at evidence from other antitrust cases, I made two proposals to help judges assess the reliability of model-based arguments. The possible extension of this framework to other areas of law where economic models are employed would be fruitful ground for future research.

31 Additionally, my suggestion aligns with the views of legal scholars such as Bartholomew (2014) and Haw Allensworth (2012), who have argued that scrutinizing the specifics of empirical models exceeds the original intent of the *Daubert* ruling. However, for a critical assessment of the gatekeeping responsibility of courts, please refer to chapter 4.

4

THE GATEKEEPER DILEMMA

Should the Daubert test apply only to the principles upon which the expert bases her testimony, or should Rule 702 also require that the application of the principles must be reliable as well? For example, with DNA tests, is it only necessary to show that the technique of DNA identification is reliable, or must it also be shown that the test was reliably conducted in the specific case?

— Daniel J. Capra[1]

4.1 INTRODUCTION

The court assessing expert testimony often finds at a critical juncture, with the weight of its decision potentially determining the course and outcome of the trial. It recognizes that, in accordance with Rule of Evidence 702, its responsibility extends beyond merely assessing the expert's qualifications and the reliability of scientific principles and methods. The court acknowledges that it must also scrutinize how these principles and methods are applied to the specific case at hand.

In the case before it, the plaintiff's economic expert introduces an intricate econometric model that lies somewhat outside the court's realm of expertise. However, a brief inquiry reveals that such a model is a common tool in the economist's toolbox and that the expert holds the title of Harvard professor. Reassured, the courts leans towards admitting the expert's testimony.

Yet, a shadow of doubt creeps in as the court revisits whether the model has been correctly applied to the circumstances of the case. Upon a careful review of the defendant's arguments, the court encounters a litany of criticisms concerning the model's underlying assumptions, the statistical techniques employed for parameter estimation, and the robustness of the analysis. Of particular concern is the omission of a variable deemed crucial by the defense, a factor they insist could entirely alter parameter estimates.

Seeking clarity on how economists decide which variables should be included in their empirical models, the court's quest reveals a landscape of varying perspectives. Some voices advocate "including all relevant variables," while others suggest the need to "select the right combination of variables." In general, the selection of appropriate explanatory variables appears to be fraught with unavoidable uncer-

1 Memorandum from Daniel J. Capra, Reporter to Advisory Comm. on Evidence Rules, to Advisory Comm. on Evidence Rules (Sept. 11, 1997), cited in Bernstein and Lasker (2015, p. 14).

tainty. The court now faces a dilemma: it can either rest its decision solely on the abstract reliability of the empirical model, thereby risking the admission of expert testimony potentially based on its misapplication, or it can delve into the specifics of the model applicability to the particular case, where certainty is elusive and judgment calls are inevitable. This the gatekeeper dilemma.

The gatekeeper dilemma will be introduced in the context of the *Daubert* standard, that is, the legal standard under which federal courts decide on the admissibility of expert witnesses. We will defend two main claims. First, courts can (and often ought to) question how scientific principles and methods have been applied to particular cases. Second, the absence of known and agreed rules of applicability in economics leads to what we refer to as the gatekeeper dilemma.[2]

In contrast to a significant portion of the existing literature, our approach rests on the assumption that judges are motivated to wield their gatekeeping power aptly in filtering scientific arguments. Additionally, we posit that, on average, judges possess the capacity to understand the arguments presented by expert witnesses. This approach allows us to trace the cause of the gatekeeper dilemma to the methodological properties of economics. Specifically, we demonstrate that the gatekeeper dilemma arises from the interaction between Rule 702, which prescribes that judges assess the correct application of scientific theories and empirical methods, and the absence of clear and agreed-upon rules for applying economic theory to specific circumstances. The result is that is exceedingly challenging for judges to evaluate the admissibility of expert testimony by economists.

The gatekeeper dilemma strikes particularly in legal areas such as antitrust disputes, where facts alone cannot speak for themselves, and the use of economic theories and models for the interpretation of those facts is *de facto* required by the law itself. The Sherman Act, the first and most important US antitrust law, refrains from precisely defining terms such as "restraint of trade" or "monopolization," implicitly deferring to external elements such as theories, models, and customs for the interpretation of the rule. The "market" concept itself is a theoretical construct, the definition of which often takes up much of the procedural history of an antitrust lawsuit and makes use of theoretical economic notions (like, say, the concepts of "cross-elasticity" and "barriers to entry"). Or consider "competition", something that antitrust law is set to protect, but whose precise meaning, and specific infringements, are entirely delegated to economic theory. It is almost inevitable, therefore, that in antitrust litigation one must face the methodological problem of what is the correct application of economic theory.[3]

2 A shorter and somewhat different version of this chapter is available here: https://philsci-archive.pitt.edu/23206/.

3 For an historical and methodological discussion on the theory-driven nature of antitrust cases, see Giocoli (2020, pp. 218–219).

4.2 THE DAUBERT STANDARD

4.2.1 *Daubert and Rule 702*

Federal courts determine the admission of expert witnesses in accordance with the Federal Rule of Evidence 702. Rule 702 is based on the precedent set by the Supreme Court in *Daubert v. Merrell Dow Pharm. Inc.* (1993), which was subsequently refined in a series of rulings referred to as the *Daubert* trilogy.[4]

Although the original *Daubert* case focused on the potential link between Mrs. Daubert's use of Bendectin and her children's serious birth defects, the Supreme Court Justices faced a broader decision: choosing whether to move beyond the existing standard for admitting scientific evidence in courts, known as the *Frye* standard.[5] The *Frye* standard originated from a 1923 murder trial where the court declined to admit a lie detector test as evidence of innocence. This standard mandated that courts assess the admissibility of scientific knowledge based on its general acceptance among the experts and practitioners in the relevant field.[6]

In contrast to the *Frye* ruling, the *Daubert* Court held that it is not enough for expert testimony to be generally accepted by the scientific community. Rather, admissibility must be based on the relevance and reliability of the evidence sought to be admitted. In the legal jargon, the *Daubert* standard has assigned courts the role of gatekeepers for the admission of scientific experts – a role they must perform by ensuring that 'any and all scientific testimony or evidence admitted is not only relevant, but reliable.' (*Daubert*, p. 589).

Expert testimony must be considered reliable only if it concerns "scientific knowledge", where the adjective "scientific" implies a "grounding in the methods and procedures of science." (*ibid*., p. 590) A few lines below, the *Daubert* Court reiterates that "the requirement that an expert's testimony pertain to 'scientific knowledge' establishes a standard of evidentiary reliability." Introducing the *Daubert* standard, therefore, represented a significant milestone in the way judges and expert witnesses interacted. Essentially, it entailed the court's determination of the admissibility of expert testimony based on its adherence to scientific methods.

4 This trilogy of cases, which occurred between 1993 and 1999, includes *Daubert v. Merrell Dow Pharm. Inc.* 509 U.S. 579, *General Electric Co. v. Joiner* 522 U.S. 136 (1997), and *Kumho Tire Co. v. Carmichael* 526 U.S. 137 (1999).

5 *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).

6 "Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while the courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made *must be sufficiently established to have gained general acceptance in the particular field in which it belongs*." (*Frye*, p. 350, emphasis added)

Finally, in 2000 the *Daubert* standard was incorporated into the new Rule 702 in the Federal Rules of Evidence. Rule 702 now reads as follows:

> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
>
> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
>
> (b) the testimony is based on sufficient facts or data;
>
> (c) the testimony is the product of reliable principles and methods; and
>
> (d) the expert has reliably applied the principles and methods to the facts of the case.[7]

The assessment of expert testimony under the *Daubert* standard typically occurs following the initiation of a *Daubert* challenge, which is a distinct legal motion presented to the judge. Parties raise these challenges before the beginning of litigation proceedings, with the aim of preventing the introduction of expert testimony to the jury. Upon receiving the challenge, the court takes the lead and assumes its gatekeeper role by applying *Daubert*/Rule 702 to evaluate the expert testimony.

On which grounds should judges evaluate whether expert testimony follows the scientific method? Justice Blackmun, who wrote for the majority in 1993, suggested four factors that courts might use when assessing whether the expert testimony is scientifically valid:

> Many considerations will bear on the inquiry, including whether the theory or technique in question can be (and has been) tested, whether it has been subjected to peer review and publication, its known or potential error rate and the existence and maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community. (*Daubert*, p. 580)

These four factors – testability, peer review, known error rate, and general acceptance – became known as the *Daubert* factors. Although the Supreme Court insisted that they did not presume to set out a definitive checklist, the *Daubert* factors are typically mentioned in most *Daubert* decisions by federal courts.

7 Retrieved from: https://www.law.cornell.edu/rules/fre/rule_702.

4.2.2 *Daubert challenges to economists*

Not surprisingly, examples of *Daubert* challenges to economists abound. Sometimes falling under the *Daubert* scrutiny are econometric techniques such as multivariate regression or instrumental variables. In *In re Live Concert Antitrust Litig.* (2012), the plaintiff's expert regression analysis was excluded by the court because it did not account for the major explanatory variables.[8] Similarly, in a 2014 *Daubert* hearing of *United States v. Am. Express Co.* (*AmEx*, for short), the US Department of Justice challenged the defendant's expert testimony based on a misapplication of the method of instrumental variables.[9]

In other cases, empirical analyses based on theoretical models from industrial economics become the target of *Daubert* challenges. Perhaps the most renowned example of a *Daubert* exclusion case involving economic experts is the one featuring Stanford economist Robert Hall. His expert testimony was dismissed in *Concord Boat Corp. v. Brunswick* (2000) due to the incompatibility of Hall's model with the economic market it aimed to represent.[10]

In support of Concord Boat's damage claim, Professor Hall relied on the Cournot duopoly model to represent the counterfactual market, that is, the market that one would observe without Brunswick's anticompetitive conduct. Applying the Cournot model, it turned out that Concord Boat and Brunswick would have shared the market equally. Despite being a standard procedure in industrial economics, the appellate judges agreed with the defendant's claim that Hall's model attributed to the challenged conduct all sales made by Brunswick above half of the total. By doing this, Hall's model ignored the fact that Brunswick already had about a 75% market share even before undertaking the challenged conduct.

In many other scenarios, the court, instead of outright rejecting the expert testimony, uses the *Daubert* hearing to carefully assess its content and validity. In a 2015 antitrust case, following a number of criticisms offered by defendants against Stanford Professor Roger Noll, a district court embarked on a careful analysis of the empirical model he employed to predict how consumers would behave if certain bundling restrictions of cable television were lifted.[11] In a lengthy *Daubert* hearing, the court explained its decision to exclude the demand side of the model while accepting the supply side.

8 863 F. Supp. 2d 966 (C.D. Cal. 2012)
9 10-CV-4496 (NGG) (RER) (E.D.N.Y. Jun. 24, 2014).
10 207 F.3d 1039, (8th Cir, 2000).
11 *Laumann v. Nat'l Hockey League* 117 F.Supp. 3d 299 (S.D.N.Y. 2015).

4.2.3 *A three-question framework*

From the perspective of philosophy of science, it seems that courts examining expert testimony by economists may raise three kinds of methodological questions:

> Are economic theories and econometric methods employed ...
>
> RQ: ... considered a reliable source of knowledge when employed by a qualified expert? (Reliability question [RQ])
>
> AQ: ... correctly applied to the particular circumstance? (Applicability question [AQ])
>
> CQ: ... capable of answering the question for which they are employed and fit with the other evidence presented at the trial? (Credibility question [CQ])

Reliability, applicability and credibility questions can take special forms in different cases. We refer to particular instances of RQ, AQ, and CQ as reliability-type questions, applicability-type questions, and credibility-type questions. For example, "Is the expert John Doe an economist?" is a reliability-type question, while "Does the regression model used by John Doe include the relevant explanatory variables?" is an applicability-type question.[12] Finally, to what extent does the regression model by John Doe demonstrate a violation of antitrust law is a credibility-type question.

There is little controversy that credibility-type questions lie outside the court's powers during a *Daubert* hearing. It is beyond the *Daubert* scrutiny to examine the model's role in determining the case outcome or its relation in the broader body of available evidence. The *Daubert* hearing is about reliability-type and applicability-type, that is, the evidence we have to "trust" the economic theories and econometric methods employed by the expert witness.

Reliability and applicability questions are asked sequentially, and the courts are rather free to raise applicability-type questions in case of a positive response to reliability-type questions. For instance, in *Concord Boat* the swing of the *Daubert* challenge was about the application of Cournot's model to represent the counterfactual market. The appellate judges explicitly raised an applicability-type question by wondering whether the assumptions of the Cournot model were realistic enough for the model to be applicable to the target market. The Court of Appeal's own words are very clear:

> Even a theory that might meet certain *Daubert* factors, such as peer review and publication, testing, known or potential error rate, and general acceptance, should not be

12 The reason for the term "applicability questions" will become evident after reading Section 4.4.

> admitted if it does not apply to the specific facts of the case. (*Concord Boat*, pp. 1056–1057)

Concord Boat best exemplifies one of the many cases where *Daubert* challenges are not about the reliability of a model or an empirical method *per se*, but rather about their application to a particular circumstance.[13] In these instances a positive response (implicit or explicit) is given to reliability-type questions, while the bulk of the *Daubert* hearing takes place on applicability.

Other cases exhibit a different attitude by federal courts. Take again the *AmEx* case, where the point of contention was whether the defendant's economic expert employed the instrumental variables method correctly. District judge Nicholas Garaufis dismissed the *Daubert* motion arguing that for what concerns the admissibility of expert testimony, the only important criterion is that regression analysis is an accepted tool in economics research and has passed muster in previous *Daubert* hearings. "Plaintiffs," the judge explained, "do not and indeed cannot, contend that multiple regression analysis is not itself a well-established and reliable econometric methodology frequently relied upon by federal courts under Rule 702." (*AmEx*, p. 8)

Using our terminology, the court refused to raise and give an answer to an applicability-type question. To be more precise, the court sidestepped the real applicability-type question (is the particular application of the instrumental variables method carried out correctly) by answering a different reliability-type question (is regression analysis a reliable method?).

To draw an analogy with the peer-review process of academic journals, the role played by reviewers tends to be mainly about applicability-type questions, while the role of journal's editors tends to be mainly about reliability-type questions. Reviewers assess the particular application of scientific methods, ensuring that they are correctly employed in the context of the research paper. They examine the details of the research's methodology and execution to determine its validity and appropriateness. On the contrary, editors take on a broader perspective by examining fundamental aspects of the manuscript, such as its overall reliability, significance, and adherence to publication guidelines. They decide whether a paper should be sent for peer review or desk-rejected based on these general considerations. According to our three-question framework, federal courts are relatively free to decide whether to assume the role of editors or reviewers toward expert testimonies.[14]

13 Further examples are provided in Peruzzi (2023). See also chapter 3, especially Section 3.6.

14 As for credibility questions, it is doubtful whether they should be a part of the peer review process. This is because including them would require a value judgment that an ideal, non-ideological referee should not impose on the argument of the article.

4.3 TWO OPEN PROBLEMS AFTER DAUBERT

4.3.1 *The paradox of expertise and court-appointed experts*

Philosophers of science have discussed the *Daubert* standard mainly from the point of view of the philosophy of expertise (e.g. Haack 2014, 2015; Martini 2015). The main concern usually takes the form of a *paradox of expertise*: how can someone who does not know the subject matter know who is an expert on that subject matter? After all, courts and judges are laypeople in comparison to the experts whose testimonies they have to evaluate; why would anyone think that they are capable of making judgments about who is and who is not an expert in a scientific field? A slightly different version of the paradox takes the following form: how can someone who does not know the subject matter know whether an expert has employed a scientific methodology?

That courts lack scientific expertise is well known to legal scholars and practitioners. In the very second ruling of the *Daubert* trilogy, Supreme Court Justice Stephen Breyer encouraged greater use of neutral experts through Rule 706, which grants judges the authority to select their own experts through the mutual agreement of the parties involved. In a discussion concerning an *amicus brief* submitted by authors from the *New England Journal of Medicine*, Justice Breyer argued that appointing judicial experts or special masters to "overcome the inherent difficulty of making determinations about complicated scientific, or otherwise technical, evidence." (*General Electric Co. v. Joiner*, 1997, p. 149).

The more frequent use of court-appointed experts aims to reduce the gap between conflicting arguments by expert witnesses and decision-makers in litigation that rely on scientific and technical evidence such as toxic tort litigation, antitrust law, employment discrimination, and more. The rationale behind the proposal is simple: if courts are clueless about scientific evidence in general, and economic evidence in particular, then we help them understand it through a court-appointed expert.

Judge Richard Posner stands out as one of the staunchest proponents of employing court-appointed experts, advocating for the use of Rule 706 in both his published articles and judicial opinions. In *High Fructose Corn Syrup Antitrust Litigation*, pp. 660–661), a class action was initiated by direct purchasers of high fructose corn syrup (HFCS) against HFCS manufacturers, with allegations of a price-fixing conspiracy.[15] The Seventh Circuit, presided over by Judge Posner, issued an opinion in 2002 where they evaluated the statistical evidence presented by the economic expert witnesses of the parties with the aim of determining the various factors that influenced the pricing of HFCS.

15 295 F.3d 651 (7th Cir. 2002).

Here is Posner's description of the battle of experts happening at the district court level:

> The plaintiffs' economic expert witness conducted a regression analysis that found, after correcting for other factors likely to influence prices of HFCS, that those prices were higher during the period of the alleged conspiracy than they were before or after. [...] The defendants presented a competing regression analysis done by one of their economic experts, who added a couple of variables to the analysis of the plaintiffs' expert and, presto, the CONSPIRE variable ceased to be statistically significant. The plaintiffs rebutted with still another expert, who pointed out correctly that adding variables that are correlated with the variable of interest can make the effect of the latter disappear – to which the defendants reply, also correctly, that there are statistical methods for solving this problem. (*High Fructose Corn Syrup Antitrust Litigation*, 2002, pp. 660–661)

Posner's conclusion on remand suggests that the district court should consider appointing its own expert. "Resolving this dispute", he says, "requires a knowledge of statistical inference that judges do not possess. Turning to the technical statistical evidence (not the data themselves, which for the most part are uncontested, but the inferences drawn from them by the use of statistical methodology), *we recommend that the district judge use the power that Rule 706 of the Federal Rules of Evidence expressly confers upon him to appoint his own expert witness*, rather than leave himself and the jury completely at the mercy of the parties' warring experts." (*ibid.*, pp. 665–666, emphasis added)

However, despite the support of many prominent scholars, the court-appointed expert option has been rarely used in *Daubert* decisions.[16] vigorously advocate for court-appointed experts, highlighting their potential role in *Daubert* hearings. "The expert", Rubinfeld and Cecil (2018) argue, "can provide useful information for a court conducting a *Daubert* hearing where the methodology or qualifications of one or more of the party-selected experts is challenged. In this setting, the technical advisor is aiding the judge in his or her role as gatekeepers.". For instance, in the realm of antitrust litigation, with nearly 300 *Daubert* hearings spanning three decades, the court-appointed expert has been invoked merely twice.[17] Practical challenges, such as the difficulty in securing a mutually agreed-upon

16 Both Sidak (2013) and Rubinfeld and Cecil (2018, pp. 157–158)

17 In *Gulf States Reorganization Grp., Inc. v. Nucor Corp.* 822 F. Supp. 2d 1201 (N.D. Ala. 2011), Judge David R. Proctor appointed a special master regarding the admissibility of plaintiff's expert testimony. Similarly, Judge Leonard P. Stark hired a special master in the antitrust class action *In re Intel Corp. Microprocessor Antitrust Litig.*, MDL Docket No. 05-1717-LPS (D. Del. Jul. 31, 2014).

expert and the parties' understandable reluctance to cover the expert's fees, likely contribute to the courts' hesitation to embrace Rule 706.[18]

4.3.2 *The problem of gatekeeping power: first and second-order analysis*

In addition to the problem of court expertise, there is another open issue that has attracted the attention of legal scholars over the past two decades: the *Daubert* ruling and Rule 702 do not sufficiently constrain the gatekeeping power of *Daubert* courts. Susan Haack characterizes the problem as follows:

> A revised FRE 702 made explicit what according to *Daubert* had been implicit in Rule 702 all along: admissible expert testimony must be based on "sufficient" data, the product of "reliable" testimony "reliably" applied to the facts of the case. Federal judges now have large responsibilities and broad discretion in screening not only scientific testimony but expert testimony generally; but they have little specific guidance about how to perform this difficult task. (Haack 2014, p. 115)

We refer to this as the *problem of gatekeeping power*, or in other words, the challenge of determining the "right amount" of gatekeeping power to be exercised in *Daubert* decisions. In a way, this matter revolves around determining the depth to which the courts should scrutinize expert testimony: excessive detail poses a risk to the authority of the jury as fact-finders, while insufficient detail represents a departure from the original intent of the *Daubert* ruling of excluding unreliable expert testimony from the trial.

To understand the right amount of gatekeeping power that courts should adopt when taking *Daubert* decisions means to answer what the judge's limits are (or should be) when evaluating expert testimony during a *Daubert* hearing. This question surely bears a legal aspect: what is the correct interpretation of the *Daubert* standard? However, it also has an epistemological aspect: is (and should) *Daubert* scrutiny be limited to checking the reliability of principles and methods in the abstract (reliability-type questions) or the application of those principles and methods to particular circumstances (applicability-type questions)? Therefore, the problem of gatekeeping power has a simple formulation in our own terms: do/should courts raise applicability-type questions?

A possible answer to such a question, based on the distinction between first and second-order analysis, has been explored in recent

18 See Lopatka (2016) for a careful analysis of the arguments against the use of the court-appointed expert. Section 4.5 demonstrates that, even setting aside these practical concerns, there are methodological reasons to doubt the added value provided by neutral experts in *Daubert* hearings.

years by legal scholars (e.g. Bartholomew 2014; Haw Allensworth 2012; Vermeule 2009). The solution contains both a descriptive and a normative part. The descriptive part argues that it is possible to distinguish between two types of gatekeeping analysis that courts can perform: first and second-order analysis. The normative part goes a step further by stating that judges should be limited, at the stage of *Daubert* hearings, to performing only second-order analysis.

At the descriptive level, the proposal goes as follows: first-order analysis would directly concern the content of the expert's testimony; in contrast to this, second-order analysis would amount to the content-neutral examination of whether the expert testimony meets certain criteria that strengthen our belief that the content of the testimony may be correct (usually referred to as *second-order criteria*). As Vermeule (2009, p. 2234) puts it:

> First-order reasons are reasons for thinking that a particular claim of fact, causation, or prediction is correct. Second-order reasons are reasons for thinking that the epistemic capacities of (a group of) experts are such that their first-order views are more or less likely to be correct.

What exactly goes into second-order analysis is a matter of considerable debate. Intuitively, second-order analysis concerns whether an expert witness has a qualified authority in the field, or whether the methods she used are usually employed for similar tasks in the relevant academic field.[19] The original *Frye* standard and some of the *Daubert* factors refer to various possible second-order criteria such as the expert's credentials, the widespread acceptance within a relevant scientific community, and whether the testimony is based on research that has passed peer-review. Second-order analysis, at least conceptually, does not go into the content of the expert testimony. In particular, it does not concern how a certain method was applied in the particular case. The analysis of the content of expert testimony and the particular application of a method constitute the domain of first-order analysis.

The distinction between first and second-order analysis aligns with the differentiation between reliability-type and applicability-type questions. Second-order analysis primarily deals with determining whether an expert witness possesses the requisite qualifications and whether the methods employed are generally accepted within the relevant field. This corresponds to addressing reliability-type questions, which assesses the reliability of the source of knowledge and the expert's authority. On the other hand, first-order analysis delves into the specific application of these methods in a particular case, addressing

19 In common law countries, second-order analysis may also assess whether similar scientific methods have already been employed in previous cases.

applicability-type questions by evaluating whether they have been correctly employed given the unique circumstances.

The normative part comes in when it is argued that the *Daubert* scrutiny should be limited to check whether the expert's methodology satisfies some agreed second-order criteria, without questioning the content of the expert testimony. In other words, courts involved in *Daubert* hearings should not ask applicability-type questions. Haw Allensworth supported this proposal in the context of economic expertise:

> Content-neutral criteria are "second order" criteria when they are used to evaluate expert opinion; in contrast, "first order" analysis would directly examine the content of the opinion. [...] In the context of economic testimony, a judge will recognize the authority of an opinion if it – and its speaker – meet certain second-order criteria set out by *Daubert* and the Federal Rules of Evidence. (Haw Allensworth 2012, p. 1278)

The rationale behind this normative proposal is clear: by limiting judicial gatekeeping to second-order analysis, it seeks to prevent courts from improperly acting as fact-finders. In other words, the goal is to preserve the different roles of judges and jurors by preventing the former from getting into the details of the content of expert testimony, with the associated risk of taking sides in the battle of experts.

4.3.3 *Challenges and pitfalls of the first and second-order analysis distinction in Daubert scrutiny*

The first/second-order divide and the associated normative proposal are intuitive and probably a fair interpretation of what the *Daubert* decision was originally meant to achieve. At the end of the day, one allegedly gets the best of both worlds: to preserve the role of jurors as a fact finder without overburdening the work of the judge, keeping her in a minimal gatekeeper role. Yet, by the aforementioned scholars' own admission, judges often fall short of this distinction and take time to examine and exclude expert testimonies after indulging in first-order analysis. "Judges," says Haw Allensworth (2012, pp. 1284–1285), "use *Daubert* to criticize the way an expert uses an accepted methodology. [...] Since *Daubert* is designed to prevent first-order analyses of expert opinions, judges must bend the standard to do so."

Finding examples that violate the normative proposal to limit courts to reliability-type questions is not difficult. In all the instances we examined in Section 4.2, federal courts have consistently posed applicability-type questions and, on occasion, opted to exclude expert testimony following unfavorable responses. Whether the issue was the

omission of explanatory variables in a regression analysis or the unrealistic nature of the assumptions in a theoretical model, economists have faced exclusion by federal courts under *Daubert* when applicability-type questions were raised.

To be fair, courts do not always pose applicability-type questions in every case. For instance, they may solely scrutinize the expert's credentials during a *Daubert* challenge, or in some instances, the court may choose not to examine the application of an econometric method in the specific case, as exemplified by *AmEx* (2014).[20]

One might naturally react to our observations by arguing that confining courts to reliability-type questions derives from normative considerations and centers on what would be optimal for courts to do, irrespective of their real-world practices. In response to this, we offer two counterarguments.

To begin, one may object that the normative proposal to limit *Daubert* decisions to second-order analysis is in fact at odds with Rule 702. As previously discussed, the existing Rules of Evidence explicitly mandate that admissible expert testimony must not only stem from reliable principles articulated by a qualified expert but also rest upon sufficient data and be reliably applied to the specifics of the case. Both the assessment of the expert's credentials and the evaluation of the reliability of methods employed fall under the umbrella of second-order analysis. Indeed, both analyses do not appear to require judges to explicitly go into the details of the content and conclusions of the testimony. However, the second part of Rule 702 suggests something more. Specifically, it asks the judge to examine the reasons – facts of the case and data – that support principles and methods used in expert testimony and whether the particular application of those principles and methods to the particular case is reliable. "Judges," as Martini (2015, p. 5) contends, "are called to judge on the reliability of principles and method of the testimony given, and on the correct application of principles and methods to the facts under investigation." To check the applicability of a given method to a particular case would usually require an examination of the content of the expert testimony, that is, some form of first-order analysis. As such, Rule 702 seems to ask the judge for more than just second-order analysis.[21]

20 In *Natchitoches Parish Hosp. Serv. Dist. v. Tyco International, Ltd.* (2009, Civil Action No. 05-12024 PBS), the motion to exclude expert testimony was based on the argument that Harvard Law Professor Einer Elhauge lacked sufficient training in economics and econometrics. This assertion was made as part of a *Daubert* challenge, with a supporting declaration from Nobel Prize winner Daniel McFadden.

21 To be sure, advocates of the first/second-order divide have recognized that by rigorously reading Rule 702, judges could end up analyzing the application of the expert's methods to the facts. "Rule 702's application prong", Bartholomew (2014, p. 2172) says, "requires courts to consider whether the economist reliably applied his methodology to the facts of the case." The requirement to limit courts to second-order analysis is tantamount to preventing judges from dealing with the application prong of Rule 702.

Our argument finds extra support in a recent analysis of the procedural history that led to the Rule 702 amendment after the *Daubert* trilogy. "The Judicial Conference", Bernstein and Lasker (2015, p. 7) explain, "made clear, or so it was thought, that courts must examine the factual basis of expert testimony and the reliability not only of expert testimony and the reliability not only of the expert's methodology but also of the application of that methodology to the facts in question."[22]

In addition to being at least questionable from a legal point of view, sticking to too rigid a distinction between first and second-order analysis is often undesirable. In many scientific areas we have generally accepted methods that, on average, generate reliable results but whose applicability to specific cases depends on various conditions that need to be checked on a case-by-case basis. Were *Daubert* scrutiny confined to second-order analysis, the risk of admitting expert testimony based on faulty applications of scientific methods would greatly increase. "At a high enough level of generality", says Hovenkamp (2005, p. 81), "virtually any methodology seems to pass muster under the *Daubert* criteria."[23]

This observation holds true for several areas of science. That regression analysis and other econometric techniques are generally accepted in the academic community, subjected to peer review, etc. (second-order criteria) tells us nothing about the validity of a particular regression model employed by an expert. In a similar vein, Haack (2014, p. 102) argues that "in the case of DNA identification, where there is solid underlying science, the most important thing is to ensure that these techniques aren't misapplied through haste, sloppiness, mismanagement, or dishonesty, conscious or unconscious." We might even think that it was exactly to avoid these potential misapplications of scientific theories and techniques that the advisory com-

22 Bernstein and Lasker propose to amend Rule 702 to make it still more explicit that "the requirement that an expert witness reliably apply his methodology to the facts of the case is essential to a trial's court gatekeeping function." (*ibid.*, p. 26) See below, Section 4.5.2.

23 It is worth quoting Hovenkamp's words in their entirety: "In too many cases the judge has observed that the expert relied on 'statistics' or 'regression analysis,' and that statistical methodologies are generally accepted in the academic community, subjected to peer review, and have a known error rate. Of course, one could say the same thing about arithmetic or geometry. At a high enough level of generality virtually any methodology seems to pass muster under the *Daubert* criteria. But statistical methodologies can be misused, and often are grossly misused. If the judge does not pass judgment on the methodology, then the job falls to the jury. Suppose, for example, that the plaintiff's expert uses statistics in a highly idiosyncratic way, perhaps making serious errors by failing to control for obvious outliers in the data, or drawing conclusions much stronger than the data permit. It is hardly useful for the judge to proclaim that 'statistics' is a widely accepted and reliable methodology of scientific investigation. Even an astrologer might use a telescope properly to observe the motions of celestial bodies, but that would not serve to validate his testimony that the alignment of the planets explained why his client murdered the victim." (Hovenkamp 2005, p. 89)

mittee, which drafted the Rules of Evidence, included the provision calling for the expert's reliable applications of principles and methods to the facts of the case.

The court's reasoning in *AmEx* epitomizes the risk of disregarding the Rule 702 application prong. Applied economists employ instrumental variables to estimate causal relationships and relies on specific assumptions whose occurrence must be checked in the particular case.[24] When these assumptions are not satisfied, the instrumental variables method provides misleading inferences about the parameter estimates and standard errors. We will return in the next section on the problems of verifying these assumptions in the application of economic models. Here, we want merely to point out that the fact that regression analysis is an accepted method and has been employed in past legal cases, provides almost no information about the particular application in the case at hand. Hence, to stop *Daubert* evaluation at second-order analysis runs the risk of admitting expert testimony based on misapplications of methods that are in the abstract reliable.

In summary, the differentiation between first and second-order analysis, along with the corresponding normative recommendation to confine *Daubert* scrutiny to the latter, is frequently contravened in judicial practice, at odds with Rule 702, and can even be undesirable in numerous instances. Therefore, we strongly advocate for the continued use of our three-question framework. This framework accurately reflects the typical behavior of courts during *Daubert* hearings, recognizing their flexibility to address both reliability-type and applicability-type questions. As demonstrated, insisting on a limitation to reliability-type questions not only contradicts the explicit language of Rule 702, but also falls short in aligning with the realities of scientific inquiry. Our framework stands as a more robust and practical approach compared to the alternative.

4.4 THE GATEKEEPER DILEMMA

Once it is established that Rule 702 allows courts to scrutinize the applicability of principles and methods to particular cases – enabling them to pose applicability-type questions – several intriguing epistemological considerations come into play. It has long been recognized that, even within well-established scientific disciplines, the effective application of theories and empirical methods hinges on the presence or absence of specific conditions. These conditions often entail the absence of interfering factors in specific applications or, equivalently, the inclusion of only those factors explicitly considered. For instance, aspirin can alleviate headaches only if certain inhibiting con-

24 The main conditions for correct estimation using instrumental variables are known as the "exclusion restriction" and the "relevance condition". See Cunningham (2021) for a textbook treatment.

ditions are absent, and a DNA test yields accurate results only when samples are properly handled and stored.

We refer to the rules specifying these conditions for successfully applying scientific theories as "rules of applicability." In practical terms, the successful application of theories and empirical methods requires adherence to the relevant rules of applicability for the specific case at hand.[25] Our three-question framework can be therefore reformulated as follows:

Are economic theories and econometric methods employed ...

RQ: ...considered a reliable source of knowledge when employed by a qualified expert?

AQ: ...applied to the particular circumstance *following the relevant rules of applicability*?

EQ: ...enable to answer the question for which they are employed and fit with the other evidence presented at the trial?

The problem, however, is that economics usually lacks rules that, once followed by the scientific expert, ensure that a theoretical model or an econometric method can be successfully applied to explain and predict a particular case. For example, there is no rule about how realistic a model's assumptions must be with respect to the target in order for it to be used to explain or predict in particular circumstances. Rules such as "model assumptions must be approximately true" or "critical assumptions must be realistic" are valid heuristic guidelines but hardly constitute an agreed-upon set of applicability rules.[26]

25 Here, we adopt the terminology found in contemporary philosophy of science, referring specifically to "scientific models" and "rules of applicability". It is worth noting, however, that the same topic has been addressed by scholars working within the logical empiricist tradition of the covering-law model. According to this tradition, to explain and predict a phenomenon means to infer it from one or more general laws and initial conditions. For example, the event "My headache will go away" is explained by deductively inferring it from a law – such as "Aspirin cures headaches" – and from a statement of initial conditions – such as "I took an aspirin." Clearly, logical empiricists have noted early on that laws and initial conditions alone are not sufficient to guarantee the occurrence of the event stated in the *explanandum*, even when all the elements of the *explanans* are true (e.g. Hempel 1965, 1988). For example, there could be some other drugs I am currently taking that inhibit the analgesic effectiveness of aspirin, thus preventing the treatment of my headache. Thus, *ceteris paribus* clauses must be introduced to state that all factors not expressed by a law remain unchanged during its application to certain phenomena to be explained or predicted. This proposal faces a well-known problem, i.e. how to make the clause more precise, so that it specifies which factors must remain equal during the application of laws for explanation and prediction. Some philosophers of economics have recognized the challenge in precisely specifying ceteris paribus clauses as a distinctive feature of economic theories in comparison to the natural sciences. (e.g. Barrotta 2000; Hausman 1992). We will reach a similar conclusion in the rest of the chapter and examine its implications for assessing economic arguments in *Daubert* hearings.

26 Rodrik (2015) defended the importance of the realism of critical assumptions when applying economic models for explanation and prediction from the perspective of

Likewise, economists do not have a rule for deciding which covariates should be included in a regression model as explanatory variables. It is well known by applied economists that the validity of our econometric models depends on the *a priori* confidence that we are considering all relevant factors in the given circumstance. As vividly pointed out by Leamer (1983, p. 32), "as every beginning econometrics student knows, if you omit from a model a variable which is correlated with included variables, bad things happen." However, there are no golden rules to understand whether a model is well specified (important variables have not been left out of the model equations) in non-experimental settings studied by economists.[27]

Those are only two simple examples, but the general lesson should be clear: in economics, rules of applicability are not easy to formulate or may not even exist. This point is not a novel one; in fact, it has a long tradition in the history of the methodology of economics. Milton Friedman, for example, makes a similar consideration about the challenge of applying economic models to particular cases:

> The rules for using the model [...] cannot possibly be abstract and complete [...]. Each occurrence has some features peculiarly its own, not covered by explicit rules. The capacity to judge that these are or are not to be disregarded, that they should or should not affect what observable phenomena are to identified with what entities in the model, is something that cannot be taught; it can be learned but only by experience and exposure in the "right" scientific atmosphere, not by rote. (Friedman 1953, p. 25)

The immediate implication for *Daubert* hearings is that judges cannot simply answer applicability-type questions by checking whether the expert testimony followed the relevant rules of applicability, because these rules themselves are open to scientific disagreement. Faced with disagreement between the two experts of the parties to the dispute, the judge cannot resort to unambiguously recognized rules of applicability of economic models and econometric methods. Courts involved

a practitioner. Philosophers of economics have explored Rodrik's proposals, eventually reaching a consensus on the usefulness of model diversity in economic research. Moreover, Rodrik and his followers seem to agree that judging which model assumptions are critical in the given application (and, thus, must be realistic) involves an undeniable subjective judgment by the economist. Indeed, as Rodrik himself reiterates in his book, "there is an unavoidable craft element involved in rendering models useful." (64). For a further analysis of Rodrik's methodological position, see the 2018 symposium of the *Journal of Economic Methodology* (Vol. 25, No. 3).

27 It is worth noting that the credibility revolution popularized by Angrist and Pischke (2010) has not been a game changer. First, the common practice in empirical research in economics is still based on taking the empirical model specification as known and rely on arbitrary decisions about the choice of explanatory variables (Moral-Benito 2015; Steel 2020). Second, in the context of legal proceedings quasi-experimental findings are hard to come by (Einav and Levin 2010).

in *Daubert* hearings about economics, therefore, found themselves in the nearly impossibility of answering the majority of applicability-type questions.

Based on the foregoing analysis, we seem to have fallen into a "gatekeeper dilemma:"

> If courts only pose reliability-type questions, then they misread Rule 702 and are likely to admit to the trial expert testimonies based on misapplications of economic theories and econometric techniques. If courts abide by Rule 702 and want to avoid the risk of admitting expert testimonies based on the misapplications of economic theories and econometric techniques, then they have to raise applicability-type questions which they cannot answer.[28]

The gatekeeper dilemma is not merely a theoretical possibility but a concrete situation where various courts have found themselves over the years. Let us go back to one of our previous examples, that is, the *AmEx* case where the defendant's economic expert employed the method of instrumental variables. In Section 4.3 we claimed that the *Daubert* decision shows the problems associated with the failure to evaluate the application of scientific methods to particular cases. While we maintain that claim, a closer analysis of the decision also illustrates the dilemma faced by the court during the *Daubert* hearing.

Intuitively, instrumental variables are used when an explanatory variable of interest is correlated with the error term, in which case ordinary least squares gives biased results. A valid instrument induces changes in the explanatory variable but has no independent effect on the dependent variable, allowing a researcher to uncover the causal effect of the explanatory variable on the dependent variable. Such a method allows for consistent estimation if and only if, among other things, the instrument is highly correlated with the endogenous explanatory variables. When the instrument is highly correlated with the endogenous explanatory variable, it is referred to as a "strong instrument". Conversely, if the correlation is low, it is termed a "weak instrument".

The core of the *Daubert* challenge in *AmEx* revolved around the allegation that the expert had used a weak instrument, implying biased estimates of parameters and standard errors. Before dismissing the challenge based on a vague appeal to the reliability of regression analysis, Judge Garaufis examined the plaintiff's arguments. In our terms, the judge proceeded to check whether the rules of applicability of the instrumental variables method had been followed by the

28 A dilemma typically involves a situation in which a hard choice must be made between two or more undesirable alternatives. The gatekeeper dilemma is such in the sense that both choices – either restricting *Daubert* scrutiny to reliability-type questions or also introducing applicability-type questions – lead us to undesirable outcomes.

defendant's expert. As a matter of fact, he quickly realized that there was no rule such that below a certain threshold the correlation between the instrument and endogenous variable is too weak for the estimation to generate unbiased results. "Plaintiffs", he claimed, "acknowledge that there exists no econometric rule dictating what particular F-statistic values demonstrate the use of weak instruments or what values necessarily render an analysis unreliable." Moreover, he added, "plaintiff's own expert has acknowledged that there exists no bright-line standard to differentiate strong from weak instruments."

The *AmEx* court faced the gatekeeper dilemma. It attempted to pose an applicability-type question, but upon realizing that it was impossible to answer due to the absence of rules of applicability, it settled for a positive response to the reliability-type question, and based on this, admitted the expert. In this way, the court chose to bear the risk of admitting into the trial an expert testimony founded on an incorrect application of the instrumental variable method.

The exact opposite happened in the *Daubert* decision by the Court of Appeals in *Concord Boat*. In response to the expert testimony, the defendant raised questions about the applicability of the Cournot model, stating that it predicts an evenly divided market only under specific assumptions – namely, when products are homogeneous and costs are uniform. "There is no dispute", the defendant argues, "that marine engines are differentiated, with manufacturers offering diverse product features, promotions, support, and [...] that engine manufacturers had different costs."[29]

Against these allegations, the plaintiff vigorously defended Dr. Hall's choice of the Cournot model:

> Dr. Hall testified [...] that the Cournot model has been shown to work quite well in predicting performance in many markets, including markets where the products are somewhat differentiated. Many eminent economists have noted the usefulness of the Cournot model. [...] Moreover, Dr. Hall extensively testified that he chose the Cournot model to use in estimating damages only *after* he conducted a comprehensive analysis of the engine market, and after having looked closely at the relationship of price to cost in this industry. (Initial Brief, p. 132, original emphasis)[30]

The court grappled with two compelling arguments. On the one hand, the defendant contended that the unrealistic assumptions underpinning Cournot's model rendered it unsuitable for accurately determining damages. On the other hand, the plaintiff maintained that

29 Reply Brief: Appellant-Petitioner. 1999 U.S. 8th Cir. Briefs LEXIS 20. Retrieved from: CONCORDBOATCORP.v.BRUNSWICKCORP.,1999U.S.8thCir.BriefsLEXIS20.

30 Initial Brief: Appellee-Respondent. 1999 U.S. 8th Cir. Briefs LEXIS 53. Retrieved from: CONCORDBOATCORP.v.BRUNSWICKCORP.,1999U.S.8thCir.BriefsLEXIS53.

economists routinely employ the Cournot duopoly model to represent markets that do not align with its foundational assumptions.

Faced with the problem of assessing the applicability of the Cournot model to the boat engine market – eminently an applicability-type question – the court could not rely on rules dictating the level of realism required for applying the model to a market. In the end, the court decided that the model's assumptions were too unrealistic for it to generate reliable conclusions.

Our main point is not whether the court made the right decision, but the fact that it was fully plunged into the gatekeeper dilemma described above. In this sense, we cannot but agree with Hovemkamp's comment on the case:

> The difficulties with the expert's model in *Brunswick* confront an old problem in the sciences: every scientific model necessarily contains fewer variables than the full set of facts that the world present. Indeed, we favor elegant models that have good predictive power while making relatively few assumptions. So *the generalist judge may have to make a judgment call about whether the expert's failure to consider some market circumstance is fatal to the model's use.* (Hovenkamp 2005, p. 83, emphasis added)

It is worth noting that the gatekeeper dilemma has nothing to do with the expertise of judges or their willingness to overrule jurors. Rather, it derives from methodological properties peculiar to economics. The dilemma would hold even if judges were motivated by the best intentions not to overtake the purview of jurors. Once they get to answer applicability-type questions and in the absence of clear and widely shared rules of applicability, judges find themselves plunged into the quicksand of the battle between experts. A comparison with another area of science frequently used in courts, such as DNA testing, will help clarify this point.

The establishment of rules of applicability plays a vital role in all scientific fields, but the clarity and consensus surrounding these rules can vary. In certain fields, a well-established and widely accepted set of rules, often referred to as "best practices," governs the application of theories and empirical techniques to specific cases.[31] For instance, DNA testing adheres to stringent guidelines to ensure accuracy, dictating factors like storage temperature and light conditions. In such cases, the gatekeeper dilemma is mitigated. Courts can efficiently evaluate the reliability of expert testimony by posing reliability-type questions and, if affirmative, proceed to examine adherence to established rules of applicability.

31 See, for example, the 2022 DOJ report on best practices for DNA identification: https://www.ojp.gov/pdffiles1/nij/304051.pdf.

In the case of evidence from DNA testing, the court could evaluate the credentials of the expert and assess the general reliability of the test employed. This evaluation involves finding answers to reliability-type questions. As a second step, the court may examine whether the rules of applicability were followed, including proper collection, preservation, and handling of the samples.

Certainly, debates concerning the rules of applicability persist across all scientific domains, and these rules are subject to modification over time. Nevertheless, we believe that the more precise, clear, and widely embraced the rules of applicability are within a scientific community, the less effort courts need to exert in assessing applicability-type questions during *Daubert* hearings. In this context, the gatekeeper dilemma we discuss is a specific challenge for economics and other disciplines characterized by the absence of clear and shared rules for applying theories and methods to particular cases, such as psychology and medicine.[32]

4.5 CONCLUSION: A WAY OUT OF THE GATEKEEPER DILEMMA?

We are not the first to point out the problems in applying the *Daubert* standard to economic theory. Indeed, numerous legal and economics scholars have repeatedly pointed out the challenges that both judges and economic experts face in the courtroom.[33] The novelty of our approach is that, rather than focusing on problems of judges' lack of scientific knowledge or the separation of the roles of court and jury, we explain that there are methodological features of economics that make *Daubert* screening inherently challenging. Is there any way out of the gatekeeper dilemma?

There are at least three solutions on the table, which represent three different ways to assist courts in making better decisions in *Daubert* hearings: a procedural one (appointing a neutral expert under Rule 706), a legal one (clarifying the meaning and scope of Rule 702 through a reform of its language), and a third educational one (bringing more economics into legal analysis). Let us examine them in order.

4.5.1 *Seeking help from neutral experts*

The proposal to appoint a neutral expert, although originally conceived to solve the problem of court expertise, may appear to be a viable solution to the gatekeeper dilemma. Court-appointed experts,

32 Incidentally, economics, psychology, and medicine are three of the four scientific areas most subjected to *Daubert* challenges. Source: www.dauberttracker.com (accessed November 17, 2023).

33 See, among others: Fisher (1980, 1986), Posner (1999), and Werden (2008). The work more closely related to our approach is Giocoli (2020).

the argument goes, may help judges taking more informed decisions about the admissibility of expert testimony. In particular, they may help court answering applicability-type questions.

Unfortunately, we believe that in the case of economic expertise and given the adversarial context, hiring a neutral expert does not provide a viable solution to the gatekeeper dilemma. A standard criticism against neutral experts, besides the question of how to choose them and who should pay for them, concerns the fact that there are no neutral experts in fields such as economics. "There is no such thing as a neutral antitrust economics expert", as Lopatka (2016, p. 456) pointed out. "Antitrust economics," he continued, "is intensely ideological. Economists begin with different preconceptions about the functioning of the markets and the likelihood of anticompetititve conduct, and those differences may have a profound impact on their analysis of the case."

Yet, we believe that court-appointed experts suffer from an even more profound problem. Courts are likely to seek the help of neutral experts to understand the validity of a particular application of a theoretical or econometric model. However, we have seen that the rules of applicability in economics are themselves open to expert disagreement. So, even a perfectly neutral expert will not be able to give truly decisive tips to the court.

Let us recall the words of the district court in *AmEx*:

> Plaintiffs acknowledge that there exists no econometric rule dictating what particular F-statistic values demonstrate the use of weak instruments or what values necessarily render an analysis unreliable. [...] Plaintiff's own expert has acknowledged that "there exists no bright-line standard to differentiate strong from weak instruments." (*AmEx*, p. 8)

Would resorting to the neutral expert have helped in such a case? We doubt it. The neutral expert, albeit not involved in party advocacy, would not have been able to provide the precise answers that Judge Garaufis was seeking. A comparable scenario arises in numerous *Daubert* disputes, especially when it concerns the incorporation of pertinent explanatory variables in a regression model. While a neutral expert could certainly offer their opinion, which might be well-founded, that the expert witness omitted crucial variables, the absence of a defined procedure for specifying the model leaves lingering doubts that the court aims to resolve.[34]

34 To be sure, we are not suggesting that court-appointed experts can have no role in a *Daubert* decision. On the contrary, the court-appointed expert surely sits in a better position than the judge to evaluate the application of both theoretical and econometric models. However, we argue that she can hardly speak the clear words that courts seek in such cases.

Eventually, another avenue exists where third-party experts can mediate between competing expert testimonies, although only in appellate cases. In American law, the *amicus curiae*, an individual or organization not involved in a case, can offer information, expertise, or insight to influence a court's decision.[35] Setting aside procedural details, we may question whether *amicus* briefs enhance the decision-making process when courts handle scientific arguments. Generally speaking, *amici curiae* resemble expert panels with clearly stated advocacy goals. Unlike expert panels that typically aim for diverse scientific perspectives, *amici curiae* gather individuals who share a specific viewpoint on the issues at hand. In this sense, all the challenges associated with the judicial evaluation of scientific arguments simply shift from opposing expert witnesses representing parties to conflicting *amici curiae* representing interest groups.[36]

An example is provided by the dueling *amicus* briefs presented in the antitrust case involving numerous states against American Express from 2010 to 2018. Without delving into excessive details, the core issue revolved around whether the credit card market should be considered a two-sided market in the assessment of antitrust violations. In February 2015, the district court sided with the Department of Justice and the states, declaring that American Express's anti-steering terms violated antitrust laws. In August 2015, four notable competition economists submitted an *amicus* brief supporting Amex's appeal.[37] Drawing on the economics literature on two-sided markets, they argued that the court failed to properly apply accepted principles in its analysis. The Appeals Court's comprehensive ruling in September 2016 overturned the lower court's decision. The states filed a petition to hear the case in June 2017, and it reached the Supreme Court. Supporting the states' petition, eight economists submitted an *amicus* brief, asserting that the Appeals Court's decision rested on an incorrect interpretation of the economics of two-sided markets.[38]

It is worth noting that in fields like antitrust and competition economics, where case characteristics can result in different theoretical interpretations, the challenges we have highlighted regarding the use of neutral experts remain. However, we maintain that examining the legal institution of *amicus curiae* from a methodological standpoint is

35 Rule 29 of the Federal Rules of Appellate Procedure. Retrieved from: https://www.law.cornell.edu/rules/frap/rule_29.

36 Unlike partisan expert witnesses, who are generously compensated, *amici curiae* operate solely out of an advocacy spirit, offering their services pro bono. However, from an intertemporal perspective, they have a vested interest in promoting their theoretical stance, as it has the potential to become established legal precedent over time.

37 *Amicus* brief available here. *Amici* were Gregory J. Sidak, Robert D. Willig, David J. Teece, and Keith N. Hylton.

38 *Amicus* brief available here. *Amici* were John M. Connor, Martin Gaynor, Daniel McFadden, Roger Noll, Jeffrey M. Perloff, Joseph A. Stiglitz, Lawrence J. White, and Ralph A. Winter.

as crucial as ever. Further exploration in this regard is warranted but falls beyond the scope of this chapter.[39]

4.5.2 *Amendments to Rule of Evidence 702*

A potential legal resolution to the gatekeeper dilemma could arise from the recent revision of Rule 702. US lawmakers have long been aware of the variety of attitudes courts take in *Daubert* hearings. Especially in recent years, there have been increasing voices pushing for reform of the language of Rule 702 precisely to clarify the court's prerogatives in the face of a *Daubert* challenge. The whole debate can be summarized in the following question: should the court limit itself to reliability-type questions or also raise applicability-type questions?

The Advisory Committee of Evidence Rules, which met in Washington, D.C. in May 2021, gave final approval for a rephrasing of the language of Rule 702 to clarify that courts can (or rather, should) ask applicability-type questions. "The Committee", we read, "unanimously favored a slight change to existing Rule 702(d) that would emphasize that the court must focus on the expert's opinion, and must find that the opinion actually proceeds from a reliable application of the methodology."[40] The Supreme Court approved the change to Rule 702, among other amendments to various federal rules, and transmitted the proposed edits to Congress in April 2023. The new rule amendments will take effect on December 1, 2023.

The amended clause (d) of Rule 702 now reads:

> Present Rule 702(d): the expert *has reliably applied* the principles and methods to the facts of the case
>
> Amended Rule 702(d): the *expert's opinion reflects a reliable application* of the principles and methods to the facts of the case.

Although the change may appear insignificant, its goal seems very clear: to halt courts from taking a passive and overly liberal role in admitting expert testimony. In our terms, the Advisory Committee aims to avoid walking the first leg of the gatekeeping dilemma, that is, having courts asking only reliability-type questions.

We look favorably on the amendment to Rule 702 because we, too, believe that there are huge risks in confining courts to reliability-type questions only. This is because, as we argued, even the most reliable methodology can lead to nonsensical results if applied incorrectly. However, we do not see the amendment as a comprehensive remedy to the gatekeeping dilemma. Certainly, in instances where rules of

39 I am grateful to one of the external reviewers of my doctoral thesis for recommending a brief examination of the *amicus curiae* case.

40 Memorandum of Advisory Committee on Evidence Rules, p. 871. Retrieved from: https://www.uscourts.gov/sites/default/files/evidence_rules_report_-_may_2022_0.pdf.

applicability are well-defined and broadly accepted, the amendment shows promise by reinforcing the imperative for the court to delve into applicability-type questions that have clear and ascertainable answers. Nevertheless, in the context of cases involving economic arguments, the amendment may fall short of providing a conclusive solution. The persistent nature of the dilemma in these situations stems from the inherent absence of clear and widely-accepted rules of applicability governing the application of economic theory to specific circumstances.

4.5.3 *A plea for the methodology of economics*

In the previous sections, we have described the existence of a gatekeeper dilemma that courts face whenever they evaluate expert opinions under the *Daubert* standard. Moreover, if we are correct, there are no quick exits from the dilemma. We want to conclude this chapter with a dreamy proposal: a defense of (perhaps better, a plea for) the methodology of economics. Gilboa et al. (2022) recently distinguished between economics, economic methods, and the methodology of economics:

> We take "economics" to include the study of various social phenomena in economics and related fields such as political science, finance, decision theory, game theory, and so forth. By "economic methods," often shortened to "methods," we refer to the development and study of techniques that economists may employ in their research. Such techniques are sometimes borrowed from other fields, including mathematics, statistics, computer science and machine learning, and are sometimes developed by economists for specific applications. The "methodology of economics," often shortened to "methodology," takes the scientific endeavor of economists as the object of enquiry. Thus, both economics and methodology belong in the social sciences, where the former deals with economic behavior, and the latter deals with the behavior of economists. Methods, by contrast, are tools that are designed to be used by scientists, but do not model a reality. (*ibid*., p. 898)

We believe that such a three-part classification system aptly characterizes the relationship between economics and its application in legal proceedings. In this context, economics serves as a scientific tool utilized by expert witnesses, encompassing both theoretical and empirical work. The methods employed by these experts are influenced by the broader scientific community, and the constant evolution of econometric techniques and modeling tools shapes the theoretical and empirical approaches taken in the courtroom. Finally, courts, tasked with

evaluating the conduct of economists and determining the admissibility of economic arguments, engage in economic methodology themselves. In light of this three-part classification, we think that any initiative aiming to enhance judges' understanding of economics (including economic theory and economic methods) must include methodological training.

Bringing more economics into legal analysis appears intuitively to be a good idea. One of the greatest challenges in antitrust cases is the increasing complexity of the economic tools used to understand markets and competition. If judges were more educated in economic theory and methods, their decisions would be more informed. Even basic microeconomic knowledge would prevent a judge from hearing the concept of oligopoly or elasticity of demand for the first time from the expert witness. Despite the consensus on the importance of educating judges in "good economics," how this should be done in practice remains rather unclear.

One notable historical experience is the Manne Economics Institute for Federal Judges – an intensive economics course attended by nearly half of US federal judges from 1976 to 1999. The Manne Institute experience is highly controversial, as it remains unclear whether judges were educated in a neutral economic theory (assuming such a theory exists) and methods, or if they were exposed to a particular perspective of economic theory associated with the so-called Chicago school. Regardless of one's perspective, the impact of the Manne Institute on the thinking of a generation of federal judges seems undeniable.[41]

While exposing courts to contemporary economic theory and methods may improve decision-making in cases involving economic evidence, we believe that equal attention should be given to raising awareness of the methodological challenges within economics.

The Manne Institute, confidentially known as "Pareto in the Pines" and later as "Pareto in the Palms" upon its move to the University of Miami in 1974, has left a lasting impact on federal judges. Inspired by this experience, we advocate for the establishment of a program – a sort of "Popper in the Pines" – aimed at educating judges in economic methodology.[42] The rationale should be clear at this point: (1) US law explicitly refers to the methodology of science through the *Daubert*

41 A recent working paper by Ash, Chen, and Naidu (2022) empirically try to quantify the impact of the Chicago-style law and economics on the US judiciary, focusing on the Manne Economics Institute for Federal Judges. Their findings indicate that judges who underwent economics training exhibited increased use of economics language in their opinions and rendered a more conservative decisions in economics-related cases. Medema and Gindis (2022) offer a historical account of Henry Manne's endeavors to institutionalize law and economics within the legal community. A good summary of the Manne Program is provided by Butler (1999).

42 Popper's choice is not accidental, considering the Popperian approach of the original *Daubert* opinion, which quotes an excerpt from *Conjectures and Refutation* (Popper 1963). Haack (2005) explains why the Popperian philosophy of science is unsuitable for aiding courts in assessing expert evidence.

trilogy; (2) the methodological status of economics is notably weak and debatable, susceptible to manipulation in a litigation-based system; and (3) therefore, there is a crucial need for judges well-versed in the fundamentals of economic methodology.

The program will center around three key questions: (i) what are economic models? (ii) how do economists apply their models to particular circumstances? (iii) what can we reasonably expect when we apply models to particular circumstances? Judges should comprehend that economists acquire knowledge about the world by constructing idealized models – simplified representations of real-world phenomena. Excluding a model under *Daubert* solely based on the existence of assumptions intended to simplify a specific market situation would constitute an error on the judge's part, misrepresenting the nature of scientific models.

Furthermore, judges should be mindful of the strategies economists employ when applying these idealized models for explanatory and predictive purposes, and acknowledge their inherent limitations. For instance, while de-idealization strategies can enhance the realism of a model by reducing the number of false assumptions, economists cannot eliminate them entirely (see chapter 2). Therefore, judges should receive education on the topic of the realism of assumptions in economics, starting from the seminal work of Friedman (1953) and the extensive debate that ensued.

Similarly, judges must understand that the selection of the "right" economic model does not guarantee its successful application to a specific case. As we have observed, the effectiveness of applying a model typically depends on numerous factors, some of which are not easily controlled (such as the quality of provided data, specific assumptions for parameter estimation, and uncertainty about functional forms).

While acknowledging that judges are already aware of the main methodological features of economics, experts may find it advantageous to be clear on the limitations of their chosen approach by explaining the nuances in applying a particular model or empirical technique to a given case. Simply put, the "Popper in the Pines" program would enhance judges' proficiency in identifying instances where experts may misuse ostensibly reliable methods to achieve predetermined outcomes.

5

CONCLUSION: LIFE BEYOND DAUBERT

Economic methodology aims to determine what constitutes a proper economic model – or, more precisely, in what sense economic models can provide valuable insights into real-world phenomena, serving various purposes such as explanation, prediction, and policy intervention. Chapter 2 has offered an example of the ongoing debate in economic methodology. Advocating for de-idealization strategies not only aligns economics more closely with natural sciences, but also provides a plausible response to skeptics who question the validity of economic models as a source of knowledge: yes, our models may contain falsehoods, but we are actively working to eliminate them.

Reflecting on the theories and methods applied in economics, along with the knowledge they yield, is far from a sterile discussion limited to armchair methodologists, but in fact can be a matter of "life or death," especially when economic models are employed for policy-making or enter a courtroom in antitrust litigation. Consider, for example, the potential harm caused by a central bank's misguided response to inflation based on a specific economic model. Such decisions can directly impact the well-being of the population. Similarly, within the realm of antitrust litigation, the acceptance of an economic model can significantly influence the trajectory of a case, potentially leading to the dismantling of a business and profound consequences for consumers and workers. Therefore, economists shoulder a tremendous responsibility when endorsing specific models, as their recommendations can have far-reaching implications for individuals and society as a whole. This is where economic methodology, which examines why economists make certain recommendations and how they justify them, takes center stage —- a theme explored in detail in both chapter 3 and chapter 4.

If I were to distill the core insights of my dissertation into two main takeaways, I propose the following. First, philosophers of economics would be wise to scrutinize the role of economic theory within the legal system, particularly in areas where economic arguments are unavoidable, such as antitrust litigation. Second, those involved in antitrust practice, alongside legal scholars and economists, would gain from revisiting some work in the methodology of economics to understand the potentials and limits when applying economic principles to particular circumstances.

My dissertation aims to serve as an initial stride toward fostering a closer connection between economists, legal scholars, and philosophers of economics. Despite my best intentions, careful readers may

have noted several omissions in the preceding chapters. To acknowledge at least some of these omissions, I will briefly address certain aspects that have been deferred for future exploration.

The first elephant in the room comes from the deliberate choice to focus on *Daubert* hearings. As mentioned in Chapter 4, the third methodological question that a court might ask to an expert witness – the enabling question – concerns whether economic theories and econometric methods are capable of answering the question for which they are employed and fit with the other evidence presented at the trial. By stating that this question is beyond the scope of the courts during *Daubert* hearings, I have refrained from an analysis of such question. In a more general sense, the question is this: once the *Daubert* hurdle is cleared, to what extent does economic theory influence the final decisions of the courts? While the issue that has occupied us so far concerns the admissibility of certain economic arguments, the enabling question pertains to the credibility (or weight) that judges or the jury attribute to these arguments in their opinions.

The answer to this question is made difficult by the absence of a metric to measure the impact of economics in the courts. Recent empirical research has confirmed the rise in economic reasoning and language in judicial decisions from various areas of law. For example, by applying computational linguistics tools to the analysis of US federal district courts' decisions from 1932 to 2016, Cao (2022) quantifies the rise of economic reasoning in court cases ranging from securities regulation to antitrust law.

Still, to say that courts are increasingly using economic language in their decisions is different from saying that economic theory has an increasingly important impact in such decisions. From a casual observation of antitrust judgments in the last twenty years, it is possible to find some variety in the courts' reactions to the use of economics. There is the contemptuous skepticism of Judge Richard Leon, who, in his opinion in the *AT&T/Warner* vertical merger case, ridiculed the bargaining model offered by the Justice Department's economic expert Carl Shapiro as a Rube Goldberg machine deprived of any factual content.[1]

There is also a different attitude, where the court acknowledges that the use of economic models has helped strengthen the plaintiff's theory and win the litigation. A case in point was the attempted merger between book publisher Penguin Random House and rival Si-

1 United States vs. AT&T Inc., 310 F. Supp. 3d 161 (DC District Court, June 2018). "After hearing Professor Shapiro's bargaining model described in open Court, I wondered on the record whether its complexity made it seem like a Rube Goldberg contraption. [...] The evidence at trial showed that Professor Shapiro's model lacks both 'reliability and factual credibility,' and thus fails to generate probative predictions of future harm." (AT&T/Warner, p. 149)

mon & Schuster.[2] The court enjoined the merger relying upon, among other things, the prediction of the likely harm to book authors arising from the merger obtained using empirical models. Despite the fact that "models are imprecise and do not perfectly reflect the way books are acquired in the publishing industry," the court concluded that "economic models generally corroborate the other evidence in the record that author advances would decrease in the wake of the merger." (*Penguin Random House*, p. 77)

Finally, there are also cases where the court found that the economic arguments provided by expert witnesses are consistent with the theory of harm and the evidence proposed by the party at trial but of limited relevance to its decision. A notable example of consistency and limited relevance was *Peabody* (2020), where the FTC wanted to block a proposed joint venture by Peabody Energy and Arch Resources, arguing it would crush competition in a region that supplies 40% of America's coal. [3] Facing a battle of experts on the appropriate modeling choices, the court washed its hands of the matter saying that "Court need not decisively sift through various models and theories." (*ibid*., p. 907)

One may hope to formally define the three attitudes with the help of elementary Bayesian probability theory. Plaintiff's economists use economic evidence mainly to reinforce the theory of harm presented by the plaintiff in the specific case, that is, the legal argument that explains how a specific business practice or behavior harms competition and violates antitrust laws. Informally, they succeed in this effort when using economic evidence increases the probability of the theory of harm in the eyes of the court adjudicating the case. In other words, an economic evidence strengthens the theory of harm as long as it gives evidential support to the theory of harm, that is, it increases the judge's subjective confidence in the theory of harm. On the contrary, economic evidence is detrimental to the theory of harm when it decreases the judge's subjective confidence in the theory of harm. An intermediate situation occurs when the use of economic evidence neither increases nor decreases the judge's subjective confidence in the theory of harm. In this sense, the economic evidence is consistent but irrelevant to the theory of harm.

Let us consider the plaintiff's theory of harm T and the available evidence E, which we assume is composed of factual, documentary, and economic evidence. For instance, that the merger will increase the prices by 5% according to a Bertrand merger simulation is part of the economic evidence available in the case. Imagine now removing the economic evidence from the available evidence E, that is, imagine a situation where the plaintiff does not hire an economic expert to

2 United States v. Bertelsmann SE & Co. KGAA, et al. WL 16949715 (D.D.C. Oct. 31, 2022). ("Penguin Random House")

3 FTC v. Peabody Energy Corp., 492 F. Supp. 3D 865 (E.d. Mo. 2020). ("Peabody").

perform economic analysis to support its case. Let us call E' the situation where the available evidence is purely factual and documentary. Thus, $P(T \mid E)$ is the probability that the theory of harm is correct given the available evidence E, while $P(T \mid E')$ is the probability that the theory of harm is correct given the available evidence E'.

Now, we can formally define the three categories of interest:

(i) Economic evidence is *consistent but irrelevant* to the theory of harm if and only if $P(T \mid E) = P(T \mid E')$.

(ii) Economic evidence is *detrimental* to the theory of harm if and only if $P(T \mid E) < P(T \mid E')$.

(iii) Economic evidence *strengthens* the theory of harm if and only if $P(T \mid E) > P(T \mid E')$.

Intuitively, we are classifying the impact of economic evidence on judicial decisions on the basis of the change in the conditional probability of the theory of harm they are purported to support. By comparing it with a situation where no economic evidence is provided, we can distinguish three cases: (i) neither increases nor decrease in conditional probability; (ii) decrease in conditional probability; (iii) increase in conditional probability.

Although this is only a sketch of a formal study of the impact of economics on judicial decisions, it suggests the possibility of addressing this issue using the theoretical framework of Bayesianism. Specifically, the three attitudes of the courts can be rationalized as the result of a variation (positive, negative, or null) in the conditional probability of a Bayesian judge.

This dissertation has a second notable gap as it exclusively deals with the American legal system, neglecting, among others, everything that happens in Europe. European competition law is shaped by articles 101 to 109 of the Treaty on the Functioning of the European Union (TFEU). The key policy areas encompass control over cartels, anticompetitive agreements, regulation of mergers, acquisitions, and joint ventures meeting specified turnover criteria, and oversight of state aid provided by EU Member States. The European Commission (EC), particularly its Directorate General for Competition, holds primary authority for enforcing competition law.

While both the European Union and the United States have competition laws aimed at promoting fair competition and preventing anticompetitive practices, there are significant differences in their legal frameworks and approaches. Werden and Froeb (2019) identify ten differences between the European and American antitrust enforcement regimes. Among the differences identified, two are particularly relevant for a methodological analysis of the role of economics in antitrust enforcement.

The first distinction is that the European system, unlike its American counterpart, operates under the purview of politicians and was originally conceived as regulation rather than a law enforcement mechanism. In a regulatory framework, the emphasis centers on establishing rules and standards for behavior, with regulatory agencies, such as the European Commission, playing a central role in overseeing and implementing these rules. As Werden and Froeb (2019) put it:

> Articles 101 and 102 of the Treaty on the Formation of the European Union use the phrase "shall be prohibited," which is legalese for "is prohibited." But that phrase also empowers an administrative agency. *The EC, rather than the courts, do the prohibiting in Europe.* (Werden and Froeb 2019, p. 3, emphasis added).

The second difference, they contend, "is that the European system does not impeach unsound theories." (*ivi*) Such a difference relates to the core theme of our discussion, namely, the presence of the *Daubert* standard that assigns to courts a gatekeeping role in excluding pseudoscience from trials:

> The US litigation system aims to screen out half-baked and dead wrong ideas in the first instance through application of the rules of evidence. [...] Since 1993 the test has been the reliability standard of *Daubert*. In the Internet Age, reliability screening is all the more important because half-baked and dead wrong ideas are so quickly and widely disseminated. Moreover, expert evidence admitted by a US court is subject to impeachment through cross examination. Nothing appears to screen out unreliable expert opinion at the EC, and nothing appears to prevent unreliable theories from being credited in EC enforcement proceedings. This makes EU enforcement more susceptible than US enforcement to political winds and passing fads. (*ibid.*, p. 5)

These distinctions serve as foundational elements for any comparison between the role of economic experts in European and American antitrust litigation and, to some extent, explain my decision to initially concentrate on the latter.

The US system prioritizes court decisions as a way to enforce antitrust laws, with its adversarial approach favoring the influence of economic experts in persuading judges. Furthermore, the existence of a standard for the admissibility of expert witnesses, albeit subject to various interpretations and not without open issues, provides a benchmark for methodologically evaluating the role of economics in legal proceedings.

Given that the purpose of this work was to assess the interaction between economic experts and the courts of justice, these two features of the US legal system provide a rational foundation for my analysis. Eventually, the availability of data on *Daubert* motions in US federal courts facilitates the empirical analysis presented in Appendix A. In any case, an analysis on the European situation would be a natural extension of the work presented here.

A

APPENDIX. AN EMPIRICAL STUDY OF DAUBERT CHALLENGES TO ECONOMISTS

While the significance of expert testimony in antitrust cases, coupled with the influential role of the *Daubert* standard, is widely acknowledged, the extent of its impact remains somewhat unclear. This appendix aims to substantiate three claims: (i) economists face more frequent challenges in antitrust litigation compared to other legal domains; (ii) the application of the *Daubert* standard places an additional onus on plaintiffs; (iii) the frequency of *Daubert* challenges exhibits an upward trend over time, thus emerging as a significant factor in numerous antitrust proceedings.[1]

A.1 DATA

The primary source of data is the online database *Daubert Tracker*, designed to comprehensively catalog all reported and numerous unreported court rulings concerning challenges to scientific experts under a variety of gatekeeping standards.[2] Representatives of the company provided upon request a customized dataset which contained court rulings on gatekeeping challenges to economic experts in all areas of law from 1993 to 2021. Notably, a substantial subset of these challenges revolves around the *Daubert* admissibility standard and Rule 702. However, the *Tracker* also gathered a set of challenges governed by alternative gatekeeping benchmarks, such as Rule 706, Rule 403, and Rule 23.[3] Therefore, not all the gatekeeping challenges included in the original dataset were made under the *Daubert* standard.[4]

1 Supplementary material, including datasets and code, can be accessed at the following link: https://drive.google.com/drive/folders/14CDFrbwESmpXLqxPi5ulDeiFII3mCEiz?usp=drive_link.

2 See: www.dauberttracker.com. I am especially grateful to Myles Levin for access to the *Daubert Tracker* database.

3 Rule 403 allows exclusion of evidence when its prejudicial impact outweighs its probative value; Rule 706 permits the appointment of expert witnesses by the court to aid the jury; Rule 23 outlines procedures for class actions in federal civil litigation. See here and here for further details.

4 As reported on their website, the *Daubert Tracker* defines a gatekeeping decision as any decision when "a) one of the cases and/or rules of evidence below has been cited or mentioned and b) a testifying expert's methodology or qualifications has been challenged." For a list of all gatekeeping challenges that the database tracked, see: https://www.dauberttracker.com/criteria.cfm.

The original dataset contains 6614 records for challenges to economic experts from 1993 to the beginning of 2021.[5] Each record also provides information on the discipline of the expert, the area of law addressed in the case (e.g. antitrust and trade regulation, patent law, product liability, etc.), the party retaining the expert, a summary of the case, a brief summary of the outcome of the challenge, and information about the parties in litigation and judges.

It is crucial to acknowledge various limitations present within the original dataset. To begin with, particularly before 2003, a significant number of records lack information about the party that retained the expert. Additionally, a more significant concern arises regarding the clarity of outcome summaries. This ambiguity sometimes makes it difficult to understand the success or failure of specific gatekeeping challenges.[6] Finally, the dataset contains numerous duplicate entries for the same challenge to an expert in the same case. Multiple entries occurs for two reasons. First, in cases where an expert's challenge in a district court leads to an appeal, multiple entries might emerge for the same expert within the same case. Second, such duplicates arise when a district court ruling references a prior gatekeeping decision within the same case.[7]

Despite these limitations, the original dataset offers a comprehensive overview of the frequency of gatekeeping challenges targeting economists across diverse legal domains. Therefore, it can be used to study whether economists have a higher probability of being challenged in antitrust compared to other legal areas. Such a study is precisely the purpose of Section A.2. It will turn out that, under reasonable assumptions about the average number of economists testifying in different legal areas, economists appearing in antitrust cases exhibit a greater probability of encountering gatekeeping challenges in comparison to their counterparts involved in alternative legal realms.

Section A.3 builds on a subset of the original dataset which contains only challenges to economists in antitrust cases. As a first step, I had selectively isolated those gatekeeping challenges that can confidently be attributed to the *Daubert* standard or Rule 702. After removing non-*Daubert* challenges, I obtained a dataset of 286 challenges to economists in antitrust cases in the period 1993-2021 which are expressly based on *Daubert* or Rule 702. I filled in missing information about the expert's name and retaining party and classified the outcome of each challenge as "Excluded," "Admissible," or "Excluded in part". More specifically, the outcome of *Daubert* challenges was re-

5 To be more precise, the original dataset contained records for all legal domains in which economists have been challenged under gateekping standards at least 50 times.

6 The dataset employs the field titled "Disposition" to indicate challenge outcomes. Regrettably, the original dataset lacks a consistent and standardized terminology for disposition descriptions.

7 In the adjusted dataset employed in Section A.3, I attempted to remove all duplicates.

viewed for each record by using data from the original dataset and searching the written court rulings on *LexisNexis*. By directly examining the court's decision, the underlying ambiguity in the "Disposition" field of the original dataset was removed.

Another data source that will be utilized below originates from the Administrative Office of the U.S. Courts. The Office has been offering statistical profiles for all 12 U.S. Courts of Appeals and 94 U.S. District Courts since 1997.[8] These profiles encompass comprehensive details concerning the annual volume of civil and criminal filings, spanning diverse legal domains. Although the U.S. Courts and *Daubert Tracker* data are not exactly comparable, an overlap exists in the classification of cases into diverse areas of law. For example, both the U.S. Courts and the *Daubert Tracker* provide the categories "Antitrust" and "Copyright, Patent, Trademark". Therefore, the yearly data from the Administrative Office will be used to compute the overall count of antitrust and patent cases handled by federal courts, thus facilitating the determination of the ratio of gatekeeping challenges to the total caseload.

A.2 WHAT IS SPECIAL ABOUT ANTITRUST LAW?

As shown in Table 1, in the period 1993-2021 there have been a total of 6614 challenges to economic experts under *Daubert* and other gatekeeping standards in the discipline of economics and its subcategories. Among them, 1166 challenges concern the area of law "Antitrust and Trade Regulation", that is, about 17.6% of all challenges. The second and third most challenged areas – "Civil Law (NEC)" and "Patent, Trademark, Copyright", which both accounts for about 12% of all challenges to economists. Therefore, in absolute terms economists have received a greater number of gatekeeping challenges in antitrust cases than in other areas of law.

Can we conclude from this result than economists testifying in antitrust cases have a higher probability of being challenged than those testifying in other legal areas? Perhaps unsurprisingly, no. A little reflection demonstrates that merely counting the number of challenges does not ensure that antitrust economists are more likely to face challenges compared to their counterparts in other legal domains.

Let us begin by noticing that the number of gatekeeping challenges x_i (simply "challenges" henceforth) to economists in the area of law i can be written as follows:

$$x_i = p_i \times n_i \times c_i \tag{28}$$

8 Refer to: https://www.uscourts.gov/statistics-reports/analysis-reports/federal-court-management-statistics.

where p_i is the probability of being challenged that an economist incurs when testifying in i, n_i is the average number of economists testifying in an i case and c_i is the total number of i cases. For example, suppose that in 10 antitrust cases, an average of two economists per case appear as expert witnesses, and the probability of facing a gatekeeping challenge is 0.3. In that case, the total number of challenges to economists would be 6.

From equation (28), we can see that the greater number of gatekeeping challenges in antitrust might be the result of two mechanisms.[9] First, on average, a greater number of economists may have testified in antitrust cases compared to other areas of law. If we assume that the probability of being challenged p_i is the same in all areas of law, then a higher average number of economists in antitrust cases means a greater number of challenges. Second, economists may have a higher probability of being challenged in antitrust cases compared to other areas of law. Suppose that the same average number of economists n_i appear in all areas of law, then a higher probability of being challenged in antitrust means a higher number of challenges.

Simply counting the number of challenges does not give us an indication on which mechanism is driving the high number of gatekeeping challenges to antitrust economists. In fact, concluding from the high number of gatekeeping challenges that antitrust economists have a higher probability of being challenged corresponds to assume away the first mechanism.

To understand the relationship between the two mechanisms, let us compare two generic legal areas, say a and b. Following (28), the total number of gatekeeping challenges to economists in the two areas is given by the following system:

$$\begin{cases} x_a = p_a \times n_a \times c_a \\ x_b = p_b \times n_b \times c_b \end{cases} \tag{29}$$

Since we are interested in p_a and p_b, we will rewrite the system as follows:

$$\begin{cases} p_a = \frac{x_a}{n_a \times c_a} \\ p_b = \frac{x_b}{n_b \times c_b} \end{cases} \tag{30}$$

p_a and p_b would give us a precise measure of the probability of being challenged that an economist incurs when testifying in the areas of law a and b. Unfortunately, however, we do not have any data for n_a and n_b. That is, we do not know the average number of economists that have appeared as expert witnesses in legal areas a and b.

9 A third mechanism, the total number of cases (c_i), can be immediately excluded because antitrust cases are only about the 0.3% of all civil cases. For comparison, patent law cases constitutes around the 4% of all civil cases.

As a second-best solution, we can study the sign of p_a and p_b for different values of n_a and n_b. To do that, let us solve the system by putting $p_a = p_b$:

$$p_a = \frac{x_a}{n_a \times c_a} = \frac{x_b}{n_b \times c_b} = p_b \tag{31}$$

from which we obtain:

$$\frac{n_a}{n_b} = \frac{x_a \times c_b}{x_b \times c_a} \tag{32}$$

Now we are ready to describe the relationship between p_a and p_b as follows:

$$\begin{cases} p_a > p_b \text{ if } \frac{n_a}{n_b} < \frac{x_a \times c_b}{x_b \times c_a} \\ p_a = p_b \text{ if } \frac{n_a}{n_b} = \frac{x_a \times c_b}{x_b \times c_a} \\ p_a \leqslant p_b \text{ if } \frac{n_a}{n_b} \geqslant \frac{x_a \times c_b}{x_b \times c_a} \end{cases} \tag{33}$$

We can also define a probability ratio $r_{a,b}$

$$r_{a,b} = \frac{p_a}{p_b} = \frac{x_a/(n_a \times c_a)}{x_b/(n_b \times c_b)} \tag{34}$$

that represents how much more likely an economist is to face a gatekeeping challenge by testifying in the legal area a rather than in b.[10] By substituting $r_{a,b}$ in the system, we obtain:

$$\begin{cases} r_{a,b} > 1 \text{ if } \frac{n_a}{n_b} < \frac{x_a \times c_b}{x_b \times c_a} \\ r_{a,b} = 1 \text{ if } \frac{n_a}{n_b} = \frac{x_a \times c_b}{x_b \times c_a} \\ r_{a,b} \leqslant 1 \text{ if } \frac{n_a}{n_b} \geqslant \frac{x_a \times c_b}{x_b \times c_a} \end{cases} \tag{35}$$

At this point, we can use the data we do have about the total number of gatekeeping challenges in both legal areas (x_a, x_b) and total cases (c_a, c_b) to study how $r_{a,b}$ varies following n_a/n_b. In what follows, we will interpret a as antitrust law and b as patent law, but the same exercise can be replicated for any two areas of law for which we have data on the total number of gatekeeping challenges and the total number of cases.

By substituting in the equation (35) the observed value of $c_b = 249060$ (total patent law cases), $c_a = 18105$ (total antitrust cases), $x_b = 779$ (total gatekeeping challenges in patent law) and $x_a = 1130$

10 For example, if $r_{a,b} = 2$, then economists testifying in legal area a are twice as likely to face a gatekeeping challenge as the ones testifying in b.

(total gatekeeping challenges in antitrust), we obtain the following system:[11]

$$\begin{cases} r_{a,b} > 1 \text{ if } \frac{n_a}{n_b} < 20 \\ r_{a,b} = 1 \text{ if } \frac{n_a}{n_b} = 20 \\ r_{a,b} \leqslant 1 \text{ if } \frac{n_a}{n_b} \geqslant 20 \end{cases} \tag{36}$$

Equation (36) tells us that, under the assumption that $n_a/n_b < 20$, economists who testify in antitrust are challenged with a higher probability than those who testify in patent law cases.[12] In other words, as long as $n_a/n_b < 20$, the higher number of gatekeeping challenges to antitrust economists cannot be driven solely by the higher average number of economists testifying in antitrust cases compared to patent law.

How realistic is that $n_a/n_b < 20$? Anecdotal evidence from legal literature indicates that economic experts frequently play a role in antitrust cases. Let us assume that, on average, two economists appeared in antitrust cases (say, one plaintiff and one defendant's expert). Our main result in (36) tells us that, as long as at least more than one economist testified on average every ten patent law cases, then $p_a > p_b$. In other words, we get $p_a < p_b$ only if we are willing to assume a very high average number of experts in each antitrust case and very few economic experts in each patent law case. For example, if it were the case that on average eight economic experts appeared in an antitrust case and only one economist for every three patent cases, then $p_a < p_b$.

Although it is realistic to think that on average there are more economists testifying in antitrust cases than in patent law cases, we do not find it credible that there are twenty times as many. For this reason, the high number of gatekeeping challenges cannot be entirely driven by the higher average number of economists testifying in antitrust cases compared to patent law cases. There seems to be something specific to antitrust cases that makes economists more likely to receive gatekeeping challenges.

A.3 DAUBERT CHALLENGES TO ECONOMISTS IN ANTITRUST CASES

The previous analysis delved into a spectrum of gatekeeping challenges raised against economic experts in federal courts. Notably, a substantial subset of these challenges revolved around the framework of *Daubert* and Rule 702. However, a cohort of challenges existed that were governed by alternative gatekeeping benchmarks, such as Rule

11 These data come from the *Daubert Tracker* and the Administrative Office of Federal Courts for the period 1997-2020.

12 We approximated the value of n_a/n_b to 20, although it would actually be 19.95.

706, Rule 403, and Rule 23. In what follows, our focus will be solely on the subset of gatekeeping challenges explicitly associated with the *Daubert* standard and Rule 702.

In addition to removing non-*Daubert* challenges, I eliminated duplicates from instances where identical *Daubert* challenges were documented multiple times within the *Daubert Tracker*. Furthermore, in cases where duplicates emerged due to appeals of *Daubert* decisions, I excluded all district-level rulings in favor of the ultimate decision reached during the appeal process. After removing non-*Daubert* challenges and duplicates, I obtained a dataset of 286 challenges to economists in antitrust cases in the period 1993-2021 which are expressly based on *Daubert* or Rule 702, covering 203 expert witnesses in 182 cases in federal courts before 174 judges.

In line with earlier empirical research on *Daubert* admissions in antitrust, Figure 3 shows that experts retained by the plaintiff are subject to significantly more challenges than those representing the defendant (Giocoli 2020; Langenfeld and Alexander 2011). More specifically, within the 286 *Daubert* challenges, plaintiff's experts account for about the 71% of all challenges (203 challenges). The remaining 29% of *Daubert* challenges were against defendant's experts (83 challenges).

Following the 286 *Daubert* challenges, economists had their opinions fully or partially excluded in 103 cases (see Figure 4). In the remaining 183 cases, the *Daubert* challenge was denied and the judge admitted the economic expert at the trial. Table 2, 3, 4, and 5 provide a more comprehensive breakdown of the outcomes of the 286 *Daubert* challenges, distinguishing between the district court and appellate reviews. Overall, 175 testimonies were deemed admissible by the district courts, 38 were excluded entirely, and 51 were partially excluded. Additionally, the appellate court reviewed the testimony in 22 cases, upholding the district court's *Daubert* decision in 18 cases, while reversing it in 4 cases.

As pointed out by Giocoli (2020), the *Daubert* standard has led to the exclusion of prominent academic economists as expert witnesses such as Dennis Carlton[13], Franklin Fisher[14], Robert Hall[15] and Robert Lucas.[16] This underscores the significant implications of *Daubert* as it extends its scrutiny to individuals with widely recognized expertise and scholarly contributions.[17]

13 *In re Delta* 2017-1 Trade Cas. (CCH) P79,944 (2017).

14 *American Booksellers Association v. Barnes & Noble, Inc.*. 135 F. Supp. 2d 1031 (2001) and *Williamson Oil Co. v. Philip Morris Cos.* 346 F.3d 1287 (11th Cir. 2003)

15 *Concord Boat Corp. vs. Brunswick Corp.* 207 F.3d 1039 (2000).

16 *In re brand Name Prescription Drugs Antitrust Litig*. 94 C 897, MDL 997 (1999).

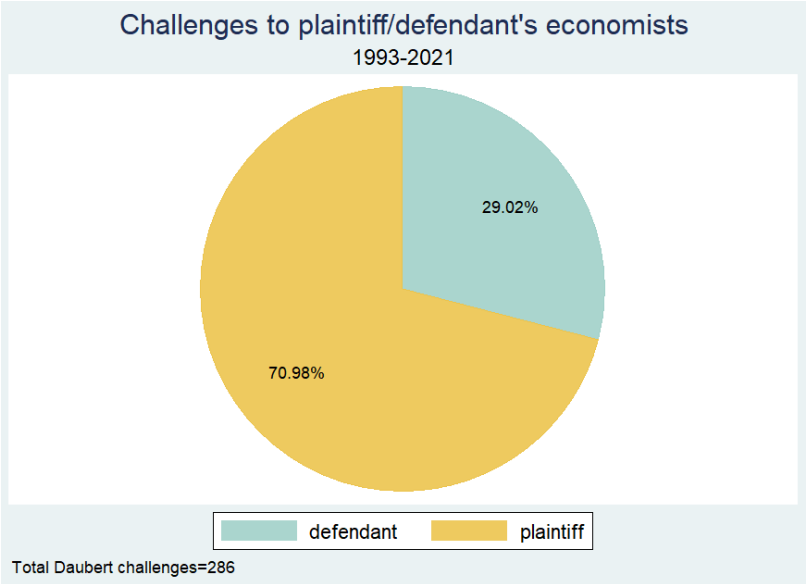
17 Additionally, other distinguished experts in industrial economics and microeconometrics, such as Robert Willig, Jerry Hausman, Orley Ashenfelter, Kevin Murphy, and Carl Shapiro, have also been the subject of *Daubert* motions.

Table 6 presents a comprehensive breakdown of the distribution of exclusion decisions across plaintiff and defendant parties. Among the 205 expert testimonies presented by the plaintiff and subjected to *Daubert* challenges, 134 were accepted as admissible, while 69 were excluded. Similarly, in the case of defendant's experts, out of 83 *Daubert* challenges, 49 were denied (and the expert admitted) and 34 were successful in excluding the expert. This contrasts with the findings of Langenfeld and Alexander (2011), revealing a higher exclusion rate for defendant's experts compared to plaintiff's experts.

Finally, let us see the trend of *Daubert* challenges over the years. Figure 5 shows both the total number of *Daubert* challenges raised every year since 1993 and the number of *Daubert* exclusions. The graph underscores the progressive rise in both *Daubert* challenges and the subsequent exclusions of expert testimonies, culminating in a pinnacle of 34 challenges during 2017.[18] Interestingly, the number of antitrust cases has not grown considerably in the same period (see Table 7 and Figure 6). Therefore, the increase in *Daubert* challenges cannot be explained by a proportional increase in the number of antitrust lawsuits. Rather, a more compelling rationale takes shape — namely, that the strategic employment of *Daubert* challenges has emerged as a prevailing practice to curtail legal disputes at an early stage. This strategic recourse has gained particular prominence within a legal domain characterized by the preeminence of expert testimony.

TABLES AND FIGURES

Figure 3: Number of *Daubert* challenges to economists by retaining party



18 The decline in the count of *Daubert* challenges and exclusions after 2020 is directly linked to the limited data coverage, which extends only to a portion of the year 2021.

Figure 4: Outcomes of *Daubert* challenges (Admissible/Excluded)

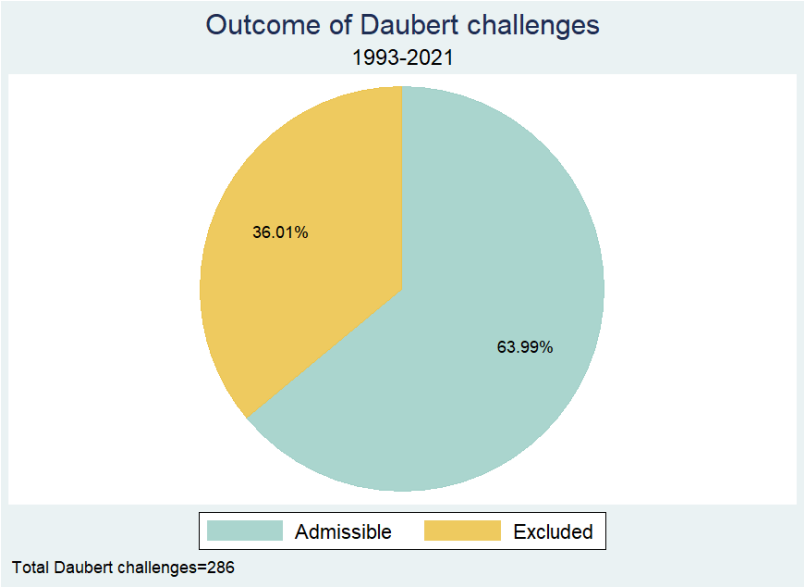


Figure 5: *Daubert* challenges and exclusions over time

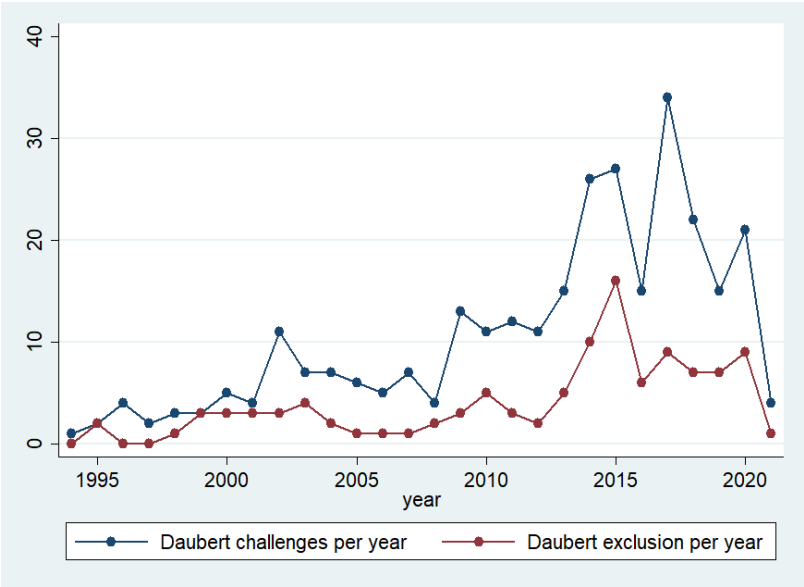


Figure 6: Antitrust cases over time

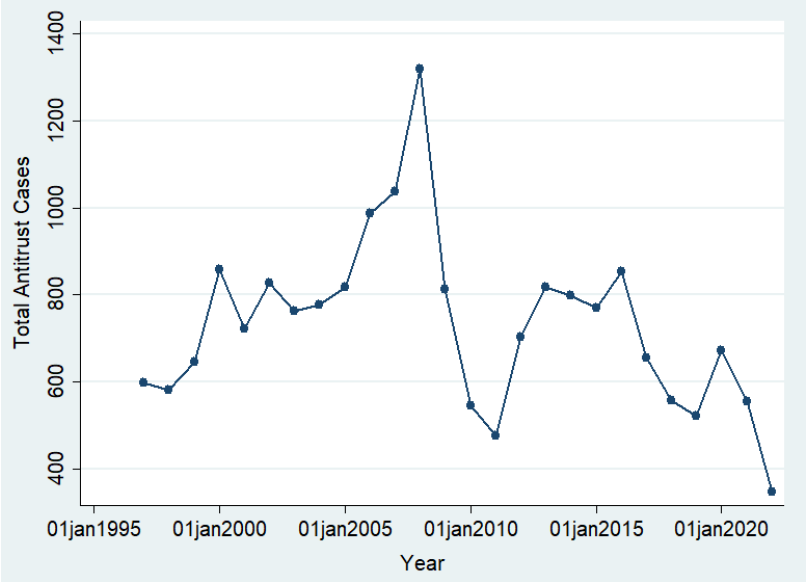


Table 1: Frequency distribution of cases by area of law

Area of Law	**Freq.**	**Percent**	**Cum.**
Antitrust & Trade Reg.	1,166	17.63%	17.63%
Bankruptcy	165	2.49%	20.12%
Civil Law (NEC)	806	12.19%	32.31%
Civil Rights	271	4.10%	36.41%
Constitutional Law	102	1.54%	37.95%
Consumer Law	127	1.92%	39.87%
Contracts	610	9.22%	49.09%
Corporate Law	59	0.89%	49.98%
Criminal Law	86	1.30%	51.29%
ERISA	51	0.77%	52.06%
Employee Benefits	52	0.79%	52.84%
Environmental Law	57	0.86%	53.70%
Fraud	76	1.15%	54.85%
Insurance Law	144	2.18%	57.03%
Labor & Employment	414	6.26%	63.29%
Medical Malpractice	106	1.60%	64.89%
Negligence	334	5.05%	69.94%
Patent, Trademark, Copyright	801	12.11%	82.05%
Personal Injury	268	4.05%	86.11%
Products Liability	206	3.11%	89.22%
Securities Law	339	5.13%	94.35%
Taxation	145	2.19%	96.54%
Tort Law (NEC)	51	0.77%	97.31%
Trade Regulation	60	0.91%	98.22%
Wrongful Death	118	1.78%	100.00%
Total	6,614	100.00%	

Table 2: Outcome of *Daubert* challenges

Outcome	Freq.	Percent	Cum.
Admissible	175	61.19	61.19
Admissible (improperly excluded by lower court)	3	1.05	62.24
Admissible (properly admitted by lower court)	5	1.75	63.99
Excluded	38	13.29	77.27
Excluded (improperly admitted by lower court)	1	0.35	77.62
Excluded (properly excluded by lower court)	10	3.50	81.12
Excluded in part	51	17.83	98.95
Excluded in part (properly excluded by lower court)	3	1.05	100.00
Total	286	100.00	

Table 3: Outcome of *Daubert* challenges in district courts

Outcome	Freq.	Percent	Cum.
Admissible	175	66.29	66.29
Excluded	38	14.39	80.68
Excluded in part	51	19.32	100.00
Total	264	100.00	

Table 4: Outcome of *Daubert* challenges in courts of appeal

Outcome	Freq.	Percent	Cum.
Admissible (improperly excluded)	3	13.64	13.64
Admissible (properly admitted)	5	23.73	33.36
Excluded (improperly admitted)	1	4.55	40.91
Excluded (properly excluded)	10	45.45	86.36
Excluded in part (properly excluded)	3	13.64	100.00
Total	22	100.00	

Table 5: Outcome of *Daubert* challenges in court of appeal

Outcome	Freq.	Percent	Cum.
Affirmed (properly admitted + properly excluded)	18	81.82	81.82
Reversed (improperly admitted + improperly excluded)	4	14.18	100
Total	22	100.00	

Table 6: Outcome of *Daubert* challenges by party

Outcome	Party		Total
	Plaintiff	Defendant	
Admissible	134	49	183
Excluded	69	34	103
Total	203	83	286

Table 7: Total cases in antitrust by year

Year	**Total Antitrust Cases**
1997	598
1998	580
1999	645
2000	858
2001	723
2002	826
2003	762
2004	777
2005	818
2006	986
2007	1038
2008	1318
2009	812
2010	544
2011	475
2012	702
2013	817
2014	799
2015	769
2016	853
2017	655
2018	557
2019	521
2020	672
2021	555

BIBLIOGRAPHY

Acosta, Juan, and Erich Pinzón-Fuchs. 2019. "Peddling Macroeconometric Modeling and Quantitative Policy Analysis: The Early Years of the SSRC's Committee on Economic Stability, 1959-1963." *Œconomia. History, Methodology, Philosophy* 9 (3): 537–558.

Alexandrova, Anna. 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, Anna. 2008. "Making models count." *Philosophy of Science* 75 (3): 383–404.

Alexandrova, Anna, and Robert Northcott. 2009. "Progress in Economics: Lessons from the Spectrum Auctions." In *Oxford Handbook for Philosophy of Economics,* edited by Harold Kincaid and Don Ross, 306–336.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of economic perspectives* 24 (2): 3–30.

Areeda, Phillip A., and Donald F. Turner. 1975. "Predatory pcricining and related practices under Section 2 of the Sherman Act." *Harvard Law Review* 88:697–733.

Ash, Elliott, Daniel L Chen, and Suresh Naidu. 2022. *Ideas Have Consequences: The Impact of Law and Economics on American Justice.* Working Paper, Working Paper Series 29788. National Bureau of Economic Research, February.

Aydinonat, Emrah N. 2018. "The Diversity of Models as a Means to Better Explanations in Economics." *Journal of Economic Methodology* 25 (3): 237–251.

Backhouse, Roger E. 2010. *The puzzle of modern economics: science or ideology?* Cambridge University Press.

Barrotta, Pierluigi. 2000. "On the role of 'ceteris paribus' clauses in economics: an epistemological approach." *History of Economic Ideas* 8 (3): 83–102.

Bartholomew, Christine P. 2014. "Death by *Daubert*: the continued attack on private antitrust." *Cardozo Law Review* 35 (6): 2147–2198.

Baye, Michael, John Morgan, and Patrick Scholten. 2006. "Information, Search, and Price Dispersion." In *Handbook on Economics and Information Systems,* edited by Terrence Hendershott, 323–375. Bingley: Emerald.

Baye, Michael R, and John Morgan. 2001. "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets." *American Economic Review* 91 (3): 454–474.

Baye, Michael R., and Joshua D. Wright. 2011. "Is antitrust too complicated for generalist judges? The impact of economic complexity and judicial training on appeals." *The Journal of Law and Economics* 54 (1): 1–24.

Belleflamme, Paul, and Martin Peitz. 2015. *Industrial Organization: Markets and Strategies.* 2nd ed. First edition: 2010. Cambridge, UK: Cambridge University Press.

Bernstein, David E., and Eric G. Lasker. 2015. "Defending daubert: it's time to amend federal rule of evidence 702." *William & Mary Law Review* 57 (1): 1–48.

Bertrand, Joseph. 1883. "Review of *Theorie mathematique de la richesse social* and of *Recherches sur les principes mathematiques de la theorie des richesses*." *Journal des Savants,* 499–508.

Björnerstedt, Jonas, and Frank Verboven. 2016. "Does merger simulation work? Evidence from the Swedish analgesics market." *American Economic Journal: Applied Economics* 8 (3): 125–64.

Bork, Robert H. 1978. *The antitrust paradox; A policy at war with itself.* New York: Free Press.

Boumans, Marcel, and Pedro Garcia Duarte. 2019. "The history of macroeconometric modeling: An introduction." *History of Political Economy* 51 (3): 391–400.

Budzinski, Oliver, and Isabel Ruhmer. 2010. "Merger simulation in competition policy: A survey." *Journal of Competition Law & Economics* 6 (2): 277–319.

Butler, Henry N. 1999. "The Manne programs in economics for federal judges." *Case Western Reserve Law Rev.* 50:351.

Cao, Siying. 2022. *Quantifying economic reasoning in court: Judge economics sophistication and pro-business orientation.* Working Paper Series 321. University of Chicago Booth School of Business.

Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement.* Oxford: Oxford University Press.

Cevolani, Gustavo, and Luca Tambolo. 2013. "Progress as Approximation to the Truth: A Defence of the Verisimilitudinarian Approach." *Erkenntnis* 78 (4): 921–935.

Chassonnery-Zaïgouche, Cléo. 2020. "How economists entered the 'numbers game': Measuring discrimination in the US courtrooms." *Journal of the History of Economic Thought* 42 (2): 229–259.

Coate, Malcolm B., and Jeffrey H. Fischer. 2012. "Daubert, Science, and Modern Game Theory: Implications for Merger Analysis." *Supreme Court Economic Review* 20 (1): 125–182.

Cools, Kees, Bert Hamminga, and Theo Kuipers. 1994. "Truth Approximation by Concretization in Capital Structure Theory." In *Idealization VI: Idealization in Economics,* edited by Bert Hamminga and Neil De Marchi, 205–228. Amsterdam: Rodopi.

Cunningham, Scott. 2021. *Causal inference: The mixtape.* New Haven & London: Yale University Press.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The rise of market power and the macroeconomic implications." *The Quarterly Journal of Economics* 135 (2): 561–644.

Dhami, Sanjit. 2016. *The Foundations of Behavioral Economic Analysis.* Oxford: Oxford University Press.

Dixit, Avinash. 1979. "A Model of Duopoly Suggesting a Theory of Entry Barriers." *Bell Journal of Economics* 10 (1): 20–32.

Dror, Itiel E. 2016. "A hierarchy of expert performance." *Journal of Applied Research in Memory and Cognition* 5 (2): 121–127.

Dror, Itiel E., and Greg Hampikian. 2011. "Subjectivity and bias in forensic DNA mixture interpretation." *Science & Justice* 51 (4): 204–208.

Eaton, B. Curtis, and Richard G. Lipsey. 1989. "Product Differentiation." In *Handbook of Industrial Organization,* edited by Richand Schmalensee and Robert Willig, 1:723–768. Amsterdam: North-Holland.

Edgeworth, Francis Y. 1925. *Papers Relating to Political Economy.* Macmillan.

Eeckhout, Jan. 2021. *The Profit Paradox: How thriving firms threaten the future of work.* Princeton, NJ: Princeton University Press.

Einav, Liran, and Jonathan Levin. 2010. "Empirical Industrial Organization: A Progress Report." *Journal of Economic Perspectives* 24 (2): 145–62.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

Finestone, Kobi. 2022. "Making Models Work." PhD diss., Duke University.

Fisher, Franklin M. 1980. "Multiple regression in legal proceedings." *Columbia Law Review* 80:702.

Fisher, Franklin M. 1986. "Statisticians, econometricians, and adversary proceedings." *Journal of the American Statistical Association* 81 (394): 277–286.

Friedman, Milton. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics*, edited by Milton Friedman, 3–43. Chicago: University of Chicago Press.

Gavil, Andrew. 2001. "Daubert comes of age." *Antitrust* 15:6–6.

Gibbard, Allan, and Hal R. Varian. 1978. "Economic Models." *Journal of Philosophy* 75 (11): 664–677.

Gilboa, Itzhak, Andrew Postlewaite, Larry Samuelson, and David Schmeidler. 2022. "Economic theory: Economics, methods, and methodology." *Revue Èconomique* 73 (2): 897–920.

Giocoli, Nicola. 2014. *Predatory Pricing in Antitrust Law and Economics: A Historical Perspective.* Routledge.

Giocoli, Nicola. 2015. "Old Lady charm: explaining the persistent appeal of Chicago antitrust." *Journal of Economic Methodology* 22 (1): 96–122.

Giocoli, Nicola. 2020. "Rejected! Antitrust Economists as Expert Witnesses in the Post-Daubert World." *Journal of the History of Economic Thought* 42 (2): 203–228.

Giocoli, Nicola. 2022. "The Chicago School and the Irrelevance of Predation." Available at SSRN: https://ssrn.com/abstract=3967604, *Competition Policy International - CPI Antitrust Chronicle* (January).

Gräbner, Claudius, and Birte Strunk. 2020. "Pluralism in Economics: Its Critiques and Their Lessons." *Journal of Economic Methodology* 27 (4): 311–329.

Grüne-Yanoff, Till. 2009. "Preface to 'Economic Models as Credible Worlds or as Isolating Tools?'" *Erkenntnis* 70 (1): 1–2.

Grüne-Yanoff, Till, and Caterina Marchionni. 2018. "Modeling model selection in model pluralism." *Journal of Economic Methodology* 25 (3): 265–275.

Haack, Susan. 2005. "Trial and error: The Supreme Court's Philosophy of Science." *American Journal of Public Health* 95 (S1): S66–S73.

Haack, Susan. 2014. *Evicence matters: science, proof, and the truth in the law.* New York: Cambridge University Press.

Haack, Susan. 2015. "The expert witness: lessons from the US experience." *Humana Mente: Journal of Philosophical Studies* 28 (2015): 39–70.

Harnay, Sophie, and Alain Marciano. 2009. "Posner, economics and the law: From "law and economics" to an economic analysis of law." *Journal of the History of Economic Thought* 31 (2): 215–232.

Hausman, Daniel M. 1992. *The inexact and separate science of economics.* Cambridge, UK: Cambridge University Press.

Haw Allensworth, Rebecca. 2012. "Adversarial economics in antitrust litigation: losing academic consensus in the battle of the experts." *Northwestern University Law Review* 106 (3): 1261–1306.

Hempel, Carl G. 1965. "The Theoretician's Dilemma: A Study in the Logic of Theory Construction." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science,* edited by Carl G Hempel, 173–226. Free Press: New York.

Hempel, Carl G. 1988. "Provisoes: A problem concerning the inferential function of scientific theories." *Erkenntnis* 28 (2): 147–164.

Hindriks, Frank. 2006. "Tractability assumptions and the Musgrave–Mäki typology." *Journal of Economic Methodology* 13 (4): 401–423.

Hindriks, Frank. 2012. "Saving truth for economics." In *Economics for Real: Uskali Mäki and the place of truth in economics,* edited by Aki Lehtinen, Jaakko Kuorikoski, and Petri Ylikoski, 43–64. London: Routledge.

Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal* 39 (153): 41–57.

Hovenkamp, Herbert. 2005. *The antitrust enterprise: Principle and execution.* Cambridge, MA: Harvard University Press.

Hovenkamp, Herbert, and Carl Shapiro. 2018. "Horizontal mergers, market structure, and burdens of proof." *The Yale Law Journal,* 1996–2025.

Jhun, Jennifer S. 2021. "Modeling the Possible to Modeling the Actual." In *The Routledge Handbook of the Philosophy of Economics,* edited by Conrad Heilmann and Julian Reiss, 316–326. Routledge.

Kirat, Thierry, and Frédéric Marty. 2021. "The late emerging consensus among American economists on antitrust laws in the 2nd New Deal (1935-1941)." *History of Economic Ideas* XXIX (1): 11–51.

Knuuttila, Tarja. 2009. "Isolating representations versus credible constructions? Economic modelling in theory and practice." *Erkenntnis* 70 (1): 59–80.

Knuuttila, Tarja, and Mary Morgan. 2019. "Deidealization: No Easy Reversals." *Philosophy of Science* 86 (4): 641–661.

Kuorikoski, Jaakko, and Aki Lehtinen. 2018. "Model selection in macroeconomics: DSGE and ad hocness." *Journal of Economic Methodology* 25 (3): 252–264.

Kuorikoski, Jaakko, Aki Lehtinen, and Caterina Marchionni. 2010. "Economic Modelling as Robustness Analysis." *British Journal for the Philosophy of Science* 61 (3): 541–567.

Kwoka, John. 2020. *Controlling mergers and market power: A program for reviving antitrust in America.* Boston, MA: Competition Policy International.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2): 443–478.

Lancieri, Filippo, Eric Posner, and Luigi Zingales. 2022. *The political economy of the decline of antitrust enforcement in the United States.* Working Paper, Working Paper Series w30326. National Bureau of Economic Research.

Langenfeld, James, and Christopher Alexander. 2011. "*Daubert* and Other Gatekeeping Challenges of Antitrust Experts." *Antitrust* 25 (3): 21–28.

Leamer, Edward E. 1983. "Let's take the con out of econometrics." *American Economic Review* 73 (1): 31–43.

Lehtinen, Aki. 2021. "The epistemic benefits of generalisation in modelling I: Systems and applicability." *Synthese* 199 (3): 10343–10370.

Lehtinen, Aki. 2022. "The epistemic benefits of generalisation in modelling II: expressive power and abstraction." *Synthese* 200 (2): 84.

Levy, Arnon. 2018. "Idealization and abstraction: refining the distinction." *Synthese*, 1–18. https://doi.org/https://doi.org/10.1007/s11229-018-1721-z.

Lisciandra, Chiara, and Johannes Korbmacher. 2021. "Multiple Models, One Explanation." *Journal of Economic Methodology* 28 (2): 186–206.

Lopatka, John E. 2016. "Economic Expert Evidence: The Understandable and the "Huh?"" *The Antitrust Bulletin* 61 (3): 434–460.

Maas, Harro, and Andrej Svorenčík. 2017. "'Fraught with controversy': organizing expertise against contingent valuation." *History of Political Economy* 49 (2): 315–345.

Mäki, Uskali. 1992. "On the Method of Isolation in Economics." In *Idealization IV: Intelligibility in Science,* edited by Craig Dilworth, 317–351. Amsterdam: Rodopi.

Mäki, Uskali. 1994. "Isolation, Idealization and Truth in Economics." In *Idealization VI: Idealization in Economics,* edited by Bert Hamminga and Neil De Marchi, 147–168. Amsterdam: Rodopi.

Mäki, Uskali. 2000. "Kinds of Assumptions and Their Truth: Shaking an Untwisted F-twist." *Kyklos* 53 (3): 317–335.

Mäki, Uskali. 2009. *The methodology of positive economics: The Milton Friedman legacy.* Cambridge, UK: Cambridge University Press.

Mäki, Uskali. 2012. "The truth of false idealizations in modeling." In *Models, Simulations and Representations,* edited by Paul Humphreys and Cyrille Imbert, 216–233. Routledge.

Mäki, Uskali. 2018. "Rights and wrongs of economic modelling: refining Rodrik." *Journal of Economic Methodology* 25 (3): 218–236.

Mäki, Uskali. 2020. "Puzzled by Idealizations and Understanding Their Functions." *Philosophy of the Social Sciences* 50 (3): 215–237.

Mandel, Michael J. 1999. "Going for the gold: Economists as expert witnesses." *Journal of Economic Perspectives* 13 (2): 113–120.

Marciano, Alain. 2023. "*Economic Analysis of Law* (1973): A methodological innovation perceived as ideological endeavor." *History of Economic Ideas* 31 (3): 33–54.

Martin, Stephen. 2010. *Industrial Organization in Context.* Oxford: Oxford University Press.

Martini, Carlo. 2015. "The paradox of proof and scientific expertise." *Humana.Mente Journal of Philosophical Studies* 28:1–16.

McMullin, Ernan. 1985. "Galilean Idealization." *Studies in History and Philosophy of Science Part A* 16 (3): 247–273.

Medema, Steven G. 2011. "Chicago price theory and Chicago law and economics: A tale of two transitions." In *Building Chicago economics: New perspectives on the history of America's most powerful economics program,* edited by Robert Van Horn, Philip Mirowski, and Thomas A Stapleford, 151–179. Cambridge: University Press Cambridge.

Medema, Steven G., and David Gindis. 2022. *One man a committee does not make: Henry manne, the AEA-AALS joint committee, and the struggle to institutionalize law and economics.* Working Paper Series No. 2022-13. Center for the History of Political Economy, December.

Milgrom, Paul, and John Roberts. 1987. "Informational asymmetries, strategic behavior, and industrial organization." *American Economic Review: Papers & Proceedings* 77 (2): 184–193.

Moral-Benito, Enrique. 2015. "Model averaging in economics: An overview." *Journal of Economic Surveys* 29 (1): 46–75.

Musgrave, Alan. 1981. "'Unreal Assumptions' in Economic Theory: The F-twist Untwisted." *Kyklos* 34:377–387.

Niiniluoto, Ilkka. 2002. "Truthlikeness and Economic Theory." In *Fact and Fiction in Economics,* edited by Uskali Mäki, 214–228. Cambridge, UK: Cambridge University Press.

Niiniluoto, Ilkka. 2012. "The Verisimilitude of Economic Models." In *Economics for Real: Uskali Mäki and the place of truth in economics,* edited by Aki Lehtinen, Jaakko Kuorikoski, and Petri Ylikoski, 65–80. London: Routledge.

Niiniluoto, Ilkka. 2018. "Explanation by Idealized Theories." *Kairos* 20 (1): 43–63.

Nowak, Leszek. 1980. *The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science.* Dordrecht: Reidel.

Paseyro Mayol, Joaquín, and Edoardo Peruzzi. 2023. "The Profit Paradox: How Thriving Firms Threaten the Future of Work, Jan Eeckhout. Princeton, NJ: Princeton University Press, 2021, viii 327 pages." *Economics & Philosophy* 39 (2): 338–343. https://doi.org/10.1017/S0266267123000044.

Peruzzi, Edoardo. 2023. "Models on trial: antitrust experts face Daubert challenges." *Journal of Economic Methodology* 30 (4): 337–351.

Peruzzi, Edoardo, and Gustavo Cevolani. 2022. "Defending De-idealization in Economic Modeling: A Case Study." *Philosophy of the Social Sciences* 52 (1-2): 25–52.

Petit, Nicolas. 2020. *Big tech and the digital economy: The moligopoly scenario.* Oxford University Press.

Popper, Karl R. 1963. *Conjectures and Refutations: the Growth of Scientific Knowledge.* London: Routledge / Kegan Paul.

Posner, Richard A. 1999. "The law and economics of the economic expert witness." *Journal of Economic Perspectives* 3 (2): 91–99.

Potochnik, Angela. 2017. *Idealization and the Aims of Science.* Chicago: University of Chicago Press.

Reiss, Julian. 2012. "Idealization and The Aims of Economics: Three Cheers for Instrumentalism." *Economics & Philosophy* 28 (3): 363–383.

Reiss, Peter C., and Frank A. Wolak. 2007. "Structural econometric modeling: Rationales and examples from industrial organization." In *Handbook of Econometrics. Volume 6A,* edited by James J. Heckman and Edward E. Leamer, 4277–4415. Amsterdam: Elsevier Sci. Ltd. North Holland.

Renault, Matthieu. 2023. "Macroeconomics under pressure: the feedback effects of economic expertise." *The European Journal of the History of Economic Thought* 30 (2): 275–298.

Rice, Collin. 2019. "Models Don't Decompose That Way: A Holistic View of Idealized Models." *British Journal for the Philosophy of Science* 70 (1): 179–208.

Rodrik, Dani. 2015. *Economics rules: the rights and wrongs of the dismal science.* Oxford: Oxford University Press.

Romer, David. 2019. *Advanced Macroeconomics.* 5th ed. First edition: 1996. New York: McGraw-Hill.

Rubinfeld, Daniel L., and Joe S. Cecil. 2018. "Scientists as experts serving the court." *Daedalus* 147 (4): 152–163.

Salop, Steven. 2021. "The FTC was correct to withdraw the vertical merger guidelines." https://www.promarket.org/2021/11/22/ftc-vertical-merger-guidelines-economics-withdrawn-lina-khan-salop/.

Shapiro, Carl. 2021a. "Antitrust: What went wrong and how to fix it." *Antitrust* 35 (3): 33–45.

Shapiro, Carl. 2021b. "Vertical mergers and input foreclosure lessons from the AT&T/Werner merger case." *Review of Industrial Organization* 59 (2): 303–341.

Shelegia, Sandro, and Chris Wilson. 2021. "A Generalized Model of Advertised Sales." *American Economic Journal: Microeconomics* 13 (1): 195–223.

Shy, Oz. 1995. *Industrial Organization: Theory and Applications.* Cambridge, MA: MIT Press.

Sidak, Gregory. 2013. "Court-appointed neutral economic experts." *Journal of Competition Law and Economics* 9 (2): 359–394.

Singh, Nirvikar, and Xavier Vives. 1984. "Price and Quantity Competition in a Differentiated Duopoly." *Rand Journal of Economics,* 546–554.

Steel, Mark F. J. 2020. "Model averaging and its use in economics." *Journal of Economic Literature* 58 (3): 644–719.

Stigler, George J. 1988. *Memoirs of an unregulated economist.* New York: Basic Books.

Sugden, Robert. 2000. "Credible Worlds: The Status of Theoretical Models in Economics." *Journal of Economic Methodology* 7 (1): 1–31.

Sugden, Robert. 2009. "Credible Worlds, Capacities and Mechanisms." *Erkenntnis* 70 (1): 3–27.

Tirole, Jean. 1988. *The theory of industrial organization.* MIT press.

Varian, Hal R. 1980. "A Model of Sales." *American Economic Review* 70 (4): 651–659.

Varian, Hal R. 1992. *Microeconomic Analysis.* 3rd ed. First edition: 1978. New York: Norton.

Vega-Redondo, Fernando. 2003. *Economics and the Theory of Games.* Cambridge, UK: Cambridge University Press.

Vermeule, Adrian. 2009. "The parliament of the experts." *Duke Law Journal* 58:2231–2275.

Weisberg, Michael. 2007. "Three Kinds of Idealization." *Journal of Philosophy* 104 (12): 639–659.

Weisberg, Michael. 2013. *Simulation and Similarity: Using Models to Understand the World.* New York: Oxford University Press.

Werden, Gregory J. 2008. "The Admissibility of Expert Testimony." In *Issues in Competition Law and Policy,* edited by American Bar Association-Section of Antitrust Law, 801–817. Aba Book Publishing.

Werden, Gregory J., and Luke M. Froeb. 2019. "Antitrust and Tech: Europe and the United States Differ, and It Matters." *Available at SSRN 3442798.*

Werden, Gregory J., Luke M. Froeb, and David T. Scheffman. 2004. "A Daubert Discipline for Merger Simulation." *Antitrust* 18:89–95.

Wimsatt, William. 2007. *Re-engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Cambridge, MA: Harvard University Press.

Wollmann, Thomas G. 2019. "Stealth consolidation: Evidence from an amendment to the Hart-Scott-Rodino Act." *American Economic Review: Insights* 1 (1): 77–94.

Woodward, James. 2006. "Some varieties of robustness." *Journal of Economic Methodology* 13 (2): 219–240.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Final Version as of May 2, 2024 (`classicthesis`).