Routledge INEM Advances in Economic Methodology

THE PHILOSOPHY OF CAUSALITY IN ECONOMICS

CAUSAL INFERENCES AND POLICY PROPOSALS

Mariusz Maziarz

The Philosophy of Causality in Economics

Approximately one in six top economic research papers draws an explicitly causal conclusion. But what do economists mean when they conclude that *A* 'causes' *B*? Does 'cause' say that we can influence *B* by intervening on *A*, or is it only a label for the correlation of variables? Do quantitative analyses of observational data followed by such causal inferences constitute sufficient grounds for guiding economic policymaking?

The Philosophy of Causality in Economics addresses these questions by analyzing the meaning of causal claims made by economists and the philosophical presuppositions underlying the research methods used. The book considers five key causal approaches: the regularity approach, probabilistic theories, counterfactual theories, mechanisms, and interventions and manipulability. Each chapter opens with a summary of literature on the relevant approach and discusses its reception among economists. The text details case studies, and goes on to examine papers which have adopted the approach in order to highlight the methods of causal inference used in contemporary economics. It analyzes the meaning of the causal claim put forward, and finally reconstructs the philosophical presuppositions accepted implicitly by economists. The strengths and limitations of each method of causal inference are also considered in the context of using the results as evidence for policymaking.

This book is essential reading to those interested in literature on the philosophy of economics, as well as the philosophy of causality and economic methodology in general.

Mariusz Maziarz is a PhD candidate at Wroclaw University of Economics, Poland, and Assistant Researcher with the Interdisciplinary Centre for Ethics & Institute of Philosophy at Jagiellonian University, Poland.

Routledge INEM Advances in Economic Methodology

Series Editor: Esther-Mirjam Sent

The University of Nijmegen, the Netherlands.

The field of economic methodology has expanded rapidly during the last few decades. This expansion has occurred in part because of changes within the discipline of economics, in part because of changes in the prevailing philosophical conception of scientific knowledge, and also because of various transformations within the wider society. Research in economic methodology now reflects not only developments in contemporary economic theory, the history of economic thought, and the philosophy of science; but it also reflects developments in science studies, historical epistemology, and social theorizing more generally. The field of economic methodology still includes the search for rules for the proper conduct of economic science, but it also covers a vast array of other subjects and accommodates a variety of different approaches to those subjects.

The objective of this series is to provide a forum for the publication of significant works in the growing field of economic methodology. Since the series defines methodology quite broadly, it will publish books on a wide range of different methodological subjects. The series is also open to a variety of different types of works: original research monographs, edited collections, as well as republication of significant earlier contributions to the methodological literature. The International Network for Economic Methodology (INEM) is proud to sponsor this important series of contributions to the methodological literature.

20 Amartya Sen and Rational Choice
The Concept of Commitment
Mark S. Peacock

21 The Nature and Method of Economic Sciences
Evidence, Causality, and Ends
Ricardo F. Crespo

22 The Philosophy of Causality in Economics
Causal Inferences and Policy Proposals
Mariusz Maziarz

For more information about this series, please visit: www.routledge.com/Routledge-INEM-Advances-in-Economic-Methodology/book-series/SE0630

The Philosophy of Causality in Economics

Causal Inferences and Policy Proposals

Mariusz Maziarz

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2020
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
52 Vanderbilt Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Maziarz, Mariusz, 1990- author.
Title: The philosophy of causality in economics : causal inferences and policy proposals / Mariusz Maziarz.
Description: Milton Park, Abingdon, Oxon ; New York, NY : Routledge, 2020. | Includes bibliographical references and index.
Identifiers: LCCN 2020002517 (print) | LCCN 2020002518 (ebook)
Subjects: LCSH: Economics—Philosophy. | Causation—Economic aspects. | Economics—Methodology.
Classification: LCC HB72 .M39 2020 (print) | LCC HB72 (ebook) | DDC 330.01—dc23
LC record available at https://lccn.loc.gov/2020002517
LC ebook record available at https://lccn.loc.gov/2020002518

ISBN: 978-0-367-36399-4 (hbk)
ISBN: 978-0-429-34642-2 (ebk)

Typeset in Bembo
by Apex CoVantage, LLC

For those motivated to read this book

Contents

Preface and acknowledgment

Did you have a drink last night? According to the recent Lancet study (Alcohol Collaborators 2018) that made headlines worldwide, the enjoyable glass of wine is a significant cause of premature death. This study is an excellent example of the difficulties connected to drawing causal conclusions from observational data. Considering this widely discussed topic can shed light on the difficulty of causal inference in economics because every second causal claim (51%, strictly speaking) is also based on a quantitative analysis of observational data (Maziarz 2018). Let me assume that you want to use the Lancet study to manipulate your wine consumption (policymaking activity). In such a case, the following three obstacles appear. First, the Alcohol Collaborators obtained their pessimistic result by including in the set of 'alcohol-related deaths' such effects that can be easily controllable by influencing their other necessary conditions. For instance, road injuries are the second most common alcohol-attributable cause of death (Alcohol Collaborators 2018, p. 9). Considering that a necessary condition for such an effect to occur is commuting while drunk, one can easily influence this factor. Restraining from driving while drunk reduces the risk of alcohol-attributable death from 25% to 33% in the 15–49 years age group (own calculation based on Alcohol Collaborators 2018, Figure 3). Second, in a related press release, the authors employed a nonstandard approach to measuring changes in relative risk depending on the number of standard drinks consumed daily. Let me discuss the change in risk of premature alcohol-related death connected to the change from zero to one standard drinks consumed on average daily. Instead of saying that drinking daily one alcoholic beverage raises the likelihood of alcohol-related premature death by 4/100,000, they decided to calculate the ratio by dividing the number of alcohol-related deaths observed among group consuming one alcoholic beverage daily into the number occurring among abstinents; i.e., 918/914. Such a procedure spuriously shows the change in likelihood to be 4% instead of 0.04%.

Apart from the problems with an unusual definition of alcohol-related health problems and nonstandard statistical procedures, the Lancet alcohol study is an excellent example of problems connected with using observational data to draw causal conclusions. Together with the widely discussed case of cigarettes causing lung cancer (e.g., Russo and Williamson 2007), they show that without either true mechanistic knowledge or randomized experiment, we cannot exclude the possibility that there is a common cause that drives both alcohol consumption and premature deaths (the so-called common-cause fallacy). For instance, people can differ in unobservable preferences regarding their health, and those who drink less may also be committed to a healthier diet (e.g., vegan, cf. Green et al. 2010) and sports lifestyle. In such a case, manipulating only the observable variable denoting average alcohol consumption may change the causal structure of the phenomena and will not lead to your longevity.

Recently, a very similar controversy connected to grounding economic policy in observational studies has stormed through economics journals. Reinhart and Rogoff (2010) and Herndon et al. (2014) arrived at inconsistent results regarding the influence of public debt on GDP growth. Both studies were used by policy-makers to justify interventions: austerity at the treasure in the former case, and following the expansion of public spending in the latter case. The divergence of results was driven mainly by choice between weighted and unweighted averaging schemes (Maziarz 2017). What these two studies have in common is hitting the headlines and being based on quantitative analyses of observational data. More mature sciences (e.g., medicine) seem to reject observational studies as insufficient evidence for drawing causal conclusions. As Grimes and Schulz (2012, p. 920) put it, "[m]ost reported associations in observational clinical research are false."

Is it implied that economists should only conduct experimental studies? If economists had unlimited resources for their research, the answer would probably be affirmative. However, due to epistemic, financial, and time limitations, economists use different evidence to draw causal conclusions. The non-experimental evidence is also useful in specific contexts. Even the results of correlational analyses can be fruitfully employed for acting in the world. For instance, insurance companies employ knowledge of car's color to estimate the likelihood of causing an accident when selling third-party liability coverage. Probably, it is not the redness of the cars that causes their drivers to drive recklessly. A common cause influencing driving style and color preferences is a more plausible explanation. In a similar vein, Reinhart and Rogoff's (2010) evidence on the relationship between public debt and economic growth can be used in a way that does not influence the causal structure. For instance, a company that considers entering a new market can choose the one least influenced by massive public debt. However, their evidence does not justify the post-crisis austerity movement (e.g., Ryan 2012) unless supported by further studies. Correlational studies can lead us astray. As Nancy Cartwright (2007, p. 33) argued:

> [t]here is a correlation between a fall in a barometer and a storm coming. However, if we manipulate the barometer in arbitrary ways (ways that vary independently from the 'other' causes of a storm), for example by smashing it, the correlation will break down.

There is no one, best research method aiming at establishing the causal structure of the world. Despite the label of 'gold standard,' even randomized field experiments (a.k.a. randomized controlled trials, or RCTs) can lead to mistaken policy decisions. For example, grounding a reform of social policy in Poland on the results of the basic income experiment conducted in Finland (cf. Koistinen and Perkiö 2014) may lead to unexpected results because of very different characteristics of the two populations (out-of-sample inference). Also, the access to alcohol can play a role.

Some economic models are liable to non-causal interpretation (Verreault-Julien 2017). Apart from causal inferences, social sciences (including economics) aim at explanation, prediction, and systematizing observations (Reiss 2007, p. 164). All these activities can be fallacious. As the Shit Academics Say tweet[1] admits, "to err repeatedly is research." However, causal claims are often directly applied to policymaking. Therefore, misuse of evidence or drawing unjustified conclusions can sometimes lead to severe consequences. Obtaining high-quality causal evidence in economics is crucial for policymaking. It is a matter of life and death. The number of suicides is related to the economic cycle in the developed countries (Weyerer and Wiedenmann 1995; Morrell et al. 1993; Leenaars et al. 1993). Furthermore, losing one's job raises the likelihood of premature death (Sullivan and Von Wachter 2009). Given that one of the primary purposes of economics is to deliver guidance for policymaking (Henschen 2018) that counteracts the inefficiencies of the economy, having accurate causal knowledge can literally save lives.

On the one hand, none of the methods of causal inference is perfect and delivers evidence that is reliable and context independent. On the other, policymakers can benefit from using even the evidence from heavily criticized observational studies if it is applied for specific purposes. Employing a causal claim put forward by economists to policymaking requires being aware of the limitations of the methods producing causal evidence and the context of research. Furthermore, understanding the meaning of the 'causal' label is crucial. The purpose of my research is to address these two questions with a view to deliver an informative guide to the methods of causal inference employed in contemporary economics and raise our understanding of economists' philosophical views on the relation between cause and effect.

This work stems from my Ph.D. research project to which I devoted myself in the autumn of 2015. Back then, when I was preparing the research proposal, I had no idea how much work was needed to complete the research and write the book. In the process, my research plans have taken their current form thanks to many people with whom I had a chance to cooperate or discuss philosophy, causality, and economics. I want to acknowledge the help and comments received from the TINT staff during my three-month research stay at the Centre for Philosophy of Social Sciences at the University of Helsinki. Special thanks go to Caterina Marchionni, Uskali Mäki, and Luis Mireles-Flores (listed alphabetically). I am also indebted to Christopher Clarke and Jack Vromen, with whom I discussed topics related to the book during my two-week research stay at the Erasmus Institute for Philosophy and Economics at Erasmus University Rotterdam. Last but not least, I need to voice my gratitude to Federica Russo and the

Institute for Logic, Language, and Computation for hosting me in November 2019 under the EPSA Fellowship scheme. The visit to Amsterdam allowed me to creatively work on the final changes to the manuscript.

Also, I need to voice my gratitude to the participants of conferences that commented on my partial results. I have presented different parts of the book at the following conferences: ENPOSS 2017, INEM 2017, EIPE 20, ENPOSS/RT 2018, the 30th Annual EAEPE Conference, and the 9th Salzburg Conference for Young Analytic Philosophy. The comments helped to improve the book. I am indebted to Julian Reiss for inviting me to present partial results at the CHESS Research Seminar and commenting on them, and all other participants (special thanks go to Nancy Cartwright and Donal Khosrowi) for delivering lots of useful comments. Furthermore, I am also grateful to the colleagues from the Polish Philosophy of Economics Network who voiced comments on numerous philosophy of economics seminars (in alphabetical order): Jarosław Boruszewski, Tomasz Dołęgowski, Marcin Gorazda, Łukasz Hardt, Paweł Kawalec, Mateusz Kucz, Tomasz Kwarciński, Robert Mróz, Krzysztof Nowak-Posadzy, and Agnieszka Wincewicz-Price). Many thanks go to my thesis advisors from Wroclaw University of Economics, Stanisław Czaja and Bartosz Scheuer, for encouragement and all kinds of help. My commitment to research would not be possible without the support of my parents and the distractions of Gabriela Staroń.

Last, but not least, I am highly indebted to the anonymous reviewers who commented on a very different earlier version of the book and my Routledge Editor, Andy Humphries. Without your trust, encouragement, guidance, and motivation, this book would not have its current shape. The research was supported by the National Science Centre, Poland (under grant no. 2015/19/N/HS1/01066). The author received a Ph.D. scholarship from the National Science Centre, Poland, under grant no. 2018/28/T/HS1/00007. All the remaining errors are author's.

Note

1 https://twitter.com/academicssay/status/596291095056617472?lang=en

References

Alcohol Collaborators (2018). Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the global burden of disease study 2016. *The Lancet*. http://dx.doi.org/10.1016/S0140-6736(18)31571-X

Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511618758

Green, L., Dare, J., & Costello, L. (2010). Veganism, health expectancy, and the communication of sustainability. *Australian Journal of Communication*, *37*(3), 51.

Grimes, D. A., & Schulz, K. F. (2012). False alarms and pseudo-epidemics: The limitations of observational epidemiology. *Obstetrics & Gynecology*, *120*(4), 920–927. DOI: 10.1097/AOG.0b013e31826af61a

Henschen, T. (2018). What is macroeconomic causality? *Journal of Economic Methodology*, *25*(1), 1–20. DOI: 10.1080/1350178X.2017.1407435

Herndon, T., Ash, M., & Pollin, R. (2014). Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. *Cambridge Journal of Economics*, *38*(2), 257–279. DOI: 10.1093/cje/bet075

Koistinen, P., & Perkiö, J. (2014). Good and bad times of social innovations: The case of universal basic income in Finland. *Basic Income Studies*, *9*(1–2), 25–57. DOI: 10.1515/bis-2014-0009

Leenaars, A. A., Yang, B., & Lester, D. (1993). The effect of domestic and economic stress on suicide rates in Canada and the United States. *Journal of Clinical Psychology*, *49*(6), 918–921. DOI: 10.1002/1097-4679(199311)49:6 < 918::AID-JCLP2270490620 > 3.0.CO;2-C

Maziarz, M. (2017). The Reinhart-Rogoff controversy as an instance of the 'emerging contrary result' phenomenon. *Journal of Economic Methodology*, *24*(3), 213–225. DOI: 10.1080/1350178X.2017.1302598

Maziarz, M. (2018). 'Emerging contrary result' phenomenon and scientific realism. *Panoeconomicus*, First View: 1–20. DOI: 10.2298/PAN171218024M

Morrell, S., Taylor, R., Quine, S., & Kerr, C. (1993). Suicide and unemployment in Australia 1907–1990. *Social Science &Medicine*, *36*(6), 749–756. DOI: 10.1016/0277-9536(93)90035-3

Reinhart, C. M., & Rogoff, K. S. (2010). Growth in a time of debt. *American Economic Review*, *100*(2), 573–578. DOI: 10.1257/aer.100.2.573

Reiss, J. (2007). Do we need mechanisms in the social sciences? *Philosophy of the Social Sciences*, *37*(2), 163–184. DOI: 10.1177/0048393107299686

Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, *21*(2), 157–170. DOI: 10.1080/02698590701498084

Ryan, P. (2012). *The Path to Prosperity: A Blueprint for American Renewal.* New York: Macmillan.

Sullivan, D., & Von Wachter, T. (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics*, *124*(3), 1265–1306. DOI: 10.1162/qjec.2009.124.3.1265

Verreault-Julien, P. (2017). Non-causal understanding with economic models: The case of general equilibrium. *Journal of Economic Methodology*, *24*(3), 297–317. DOI: 10.1080/1350178X.2017.1335424

Weyerer, S., & Wiedenmann, A. (1995). Economic factors and the rates of suicide in Germany between 1881 and 1989. *Psychological Reports*, *76*(3), 1331–1341. DOI: 10.2466/pr0.1995.76.3c.1331

1 Introduction

Despite causal talk in economics regaining in popularity nowadays (Hoover 2004), philosophy of economics lacks a systematic study of the methods of causal inference. The majority of philosophers interested in causal inferences in economics attempt at guiding economists and conduct normative analyses. For instance, Kevin Hoover (2001) developed a technique (known as Hoover's test) for solving the problem of determining the direction of causal relations using knowledge of changes in policymaking, and Tobias Henschen (2018) argued for the inconclusiveness of causal evidence in macroeconomics. The hitherto descriptive research on the topic is fragmentary. Four different approaches can be distinguished. First, some analyses focus on discussing the points of view on causation presented by philosophers of economics. Second, studies focus on the historical development of a chosen method or analyze the approach to causal inference practiced by a famous economist. Third, some philosophers attempt to review the philosophy of causality literature and making it relevant to economic research. Finally, a few notable studies focus on analyzing chosen cases of causal economic research with a view to uncovering the meaning of causality presupposed by economists. However, because of a limited sample, they do not deliver a systematic knowledge of causal inferences in economics. Francois Claveau and Luis Mireles-Flores studied the meaning of causal generalizations employing referentialist (2014) and inferentialist (2016) semantics using an OECD report on unemployment as an example. On the grounds of a few case studies, Tobias Henschen (2018) supported a manipulationist definition of causality as adequate to macroeconomics.

Today, some questions connected to causal inferences in economics stay open. On what grounds causal conclusions are put forward, what 'causality' means for economists, and what philosophical assumptions underlie the methods used for causal inference are the problems I address in the book. My research aims at developing our understanding of causal research in economics and helping policymakers understand the limitations of employing causal conclusions to intervening in the world of economy. On the one hand, I want to develop the philosophy of economics literature by analyzing the methods of causal inference with the view to reconstructing the meaning of causal claims put forward by economists and philosophical presuppositions underlying the research methods they use. On the other, I strive for writing a guide to research methods employed by economists that will prove useful for economists and policymakers. In this chapter, I want to review the philosophical debate on causality (Section 1.1), describe the method of inquiry used in my research (Section 1.2), and summarize the topics covered in each chapter of the book (Section 1.3).

1.1 The meaning of causality

The topic of causality seems to be as old as philosophy itself. Despite over two-millennia-long discussions, there is little consensus concerning the question of what causality is. Aristotle coined the first fully fledged theory of causality (cf. Lossee 2012; Reiss 2015, p. 2). Among the four types of causes (formal, material, efficient, and final) distinguished by Aristotle, only the efficient one is of interest for the contemporary science. Today, the 'why' questions denote interest in the factors driving a phenomenon seeking an explanation. The contemporary stance contradicts the viewpoint of the Greek scholar, who believed that "[i]t is the job of the natural scientist, then, to understand all four of these causes; if he refers to the question 'Why?' to this set of four causes – matter, form, source of change, purpose" (Aristotle 1999, p. 49).

Contrary to Aristotle, who coined a philosophical theory of causality, Robert Grosseteste, an eleventh-century scholastic philosopher and scientist, entered the pages of history-of-science books as being interested in the practice of causal inferences. Attempting at discovering which herbs cause demanded effects, he coined an inductive-reasoning procedure (Grosseteste's procedure) aimed at indicating a difference-making factor (Serene 1979; Lossee 2012, pp. 7–9). This method resembles John Stuart Mill's (1893) method of difference.

Another essential step in the process of developing our views on what causation is and how such relations are inferred is the exclusion of the Aristotelian final causes from scientific inquiry in the early seventeenth century (Lossee 2012, pp. 11–13). As Francis Bacon (1852, p. 47) put it,

> [b]y a 'final cause' . . . is meant the use or end which the Creator had in view in establishing this or that arrangement. Such enquiries are very far indeed from being unprofitable; but, by the followers of Aristotle, they were perhaps too much mixed up with the enquiry regarding physical causes.

Another step in the process of returning from metaphysical considerations was taken when Cartesian philosophers criticized the Newtonian law of gravitational attraction for accepting "action at a distance" (cf. Lakatos 1980, p. 47). Despite the later acceptance of Newton's physics, the viewpoint according to which the relationship between cause and effect is to be spatiotemporally united was crucial for the development of the Humean viewpoint on science (Chapter 2). Although David Hume is recognized as the father of the reductionist approach to causality (Tooley 1990), the Scottish philosopher is eligible for various interpretations due to internal inconsistencies (Converty 2006).

Nevertheless, he usually gets the credit for starting the modern philosophical debate on causality.

The contemporary debates have led to forming several opposing views on nature (ontology) and appearances (epistemology) of causality. The voices raised in the contemporary debate can be classified into several 'families of theories' or 'approaches' to causality. The standard stance (Beebee et al. 2012; Reiss 2012) divides the literature into five main approaches. Namely, they are the (1) regularity, (2) probabilistic, (3) counterfactual, (4) causal-process (mechanistic),[1] and (5) manipulationist approaches to causality. Roughly speaking, causal relations are defined in terms of (1) empirical, observable regularities, (2) changes in conditional probabilities, (3) difference-making conditions, (4) mechanisms underlying relations, and (5) invariance under intervention. Despite these generalizing descriptions, each approach to causality is internally differentiated and consists of numerous philosophical theories of causality. I discuss each of these philosophical approaches in a separate chapter of the book (cf. Section 1.3) and relate them to different methods of causal inference practiced in economics.

Some philosophers believe that all the standard theories of causality are reductive. They attempt to define causality by reducing such relations to different kinds of relations such as manipulability, co-existence, or correlation (cf. Carroll 2009). This criticism is not justified when one considers the distinction between ontic and epistemic dimensions. Some of the theories belonging to the main five approaches to causality are not reductionist but, instead, they focus on how causal relations can be observed (researched) in the world. In fact, all main five approaches to causality can be interpreted as epistemic stances. An excellent example of such an approach is the Humean distinction between the regularity (appearance) and counterfactual (nature) view on cause and effect.

James Woodward (2015) acknowledged that the philosophy of causality could be divided into the following two types of studies. On the one hand, philosophers investigate what causality is (ontology of causality) and how can we infer causal relations (epistemology of causality). On the other hand, philosophers and psychologists address the question of how different groups (adults, children, physicists, etc.) understand the relation between cause and effect. As Claveau and Mireles-Flores (2014, p. 403) put it, "[i]n economics, it is seldom explicit which philosophical theory of causality is supposed to hold."

1.2 On referentialist semantics, case studies, and the choice of sample

The use of causal label by economists implicates that either economists interpret their study as delivering evidence for the presence of this type of relation or use this name with a view to underline the importance of conclusions and their application to policymaking. In either case, the question of what 'causality' means arises. The use of specific research methods can give a hint of which of the five approaches of causality is implicitly presupposed by economists. Alternatively, my analysis focusing on studying the research methods used by economists to draw causal conclusions aims at uncovering the meaning of causal conclusions that can justifiably be drawn from each method of causal inference employed by economists.

Semantics is the branch of philosophy that studies meaning in a formal way. The mainstream stance in this field today is the approach labeled 'referentialist semantics' (Bianchi 2015, p. 2) that dates back to Gottlob Frege (1948 [1892]). According to this stance, the meaning of expressions is given by their extensions (referents); i.e., the objects to which they refer. On the one hand, this idea that the meaning of the sentence 'the sun shines' is given by the shining sun seems obvious but, on the other, the following problem occurs: since direct comparison of linguistic expressions and real-world entities is impossible, talking about the extensions of expressions in a meaningful way is impossible. Putting the philosophical problems and criticism formulated within the semantics literature aside, the referentialist semantics offer a useful hint as to how the meaning of words in use can be reconstructed. In this book, I want to analyze 'semantic reference,' i.e., the relation between a word and things in the world that it describes (Gauker 2015, p. 22), of the causal-family terms used by economists to conclude the causal inferences. For simplicity, let me discuss the following thought experiment. A person with limited knowledge of English asks "where is 'akuulakärkikynä'"? Considering that you do not know Finnish, you cannot guess what they are looking for. However, if your colleague finally finds his pen and starts writing, you can suppose with some degree of certitude that 'akuulakärkikynä' refers to the same group of objects that the English 'pen' does.

By analogy, I attempt at understanding how economists understand causality by analyzing what the causal words refer to. Let me consider the example of an economist concluding that 'high levels of public debt hamper economic growth.' What they have in mind is that when public debt exceeds the level believed to be 'high,' then it 'causes' slower economic growth. The economy is an extremely complex phenomenon. Its analysis is only possible by means of research methods and models aimed at idealizing (usually simplifying) or isolating the aspects of economic reality chosen by economists. For these purposes, economists use various theoretical and empirical models, and study economy by analyzing their models.[2] For example, an economist interested in the influence of public debt on GDP (Figure 1.1) employs a research method with a view to depicting the relation under consideration (bold arrow) and excludes other factors (regular arrows) from the analysis. On the grounds of the study, they conclude that public debt influences GDP. Considering that economists use models (*sensu largo*) because they do not have direct access to the economy, analyzing the relation of reference between causal claims and the economic reality in a meaningful way is impossible. Therefore, I believe that the causal claims put forth by economists refer to the results produced by the use of research methods (dotted lines). In other words, the question if the research methods are successful at depicting causal structure of the world falls beyond the scope of my research.

Economists who conclude their research with 'causes,' 'causal,' 'drives,' 'influences,' 'lowers,' 'raises,' and similar causal-family words are likely to refer to the

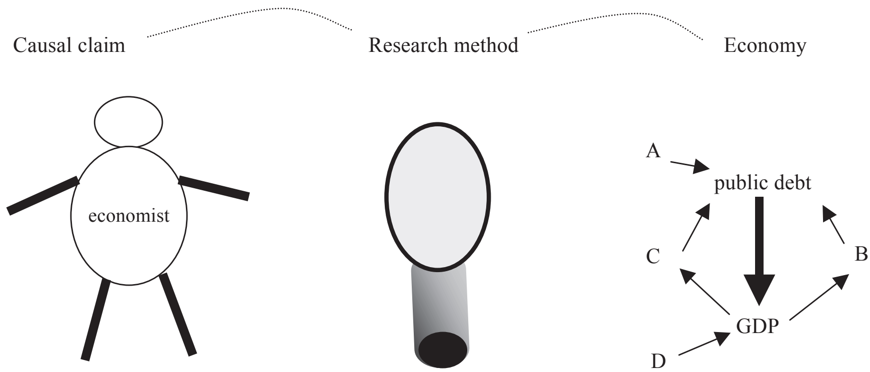

Figure 1.1 The referents of causal claims

results of their studies. In other words, the causal words refer to what can be discovered with the research methods they use. For example, if a researcher calculates the correlation between *A* and *B* and concludes that *A* causes *B*, the causal label is likely to refer to that correlation. Similarly, the use of randomized controlled trial to establish that customers buy less when the prices are higher suggests that 'lower' in the conclusion 'higher prices lower demand' refers to the following manipulationist-neutral condition: raising (lowering) prices raises (lowers) demand.

Two problems occur. First, economists can implicitly accept only one definition of causality but use various research methods to analyze causal relations. For example, the use of experiments does not exclude the possibility that the experimenter defines causality in terms of constant conjunction of two events. Therefore, to limit the indeterminacy of views on causality, I assume that economists employ the most cost-effective research methods. In other words, I try to interpret economists' views on causality in a most epistemically courageous way, having in mind that economics is a policy-oriented science and causal claims should have direct application to policymaking. Therefore, my analysis can serve as a guide for policymakers who want to use causal evidence to intervene in the economy. Second, the distinction between ontic and epistemic meanings (definitions) of causality raises a similar obstacle. Ontology (also known as metaphysics) is a branch of philosophy that investigates what exists and the features (nature) of entities. Epistemology, roughly speaking, addresses the questions of what can be known and how we can justify our knowledge. Ontic and epistemic definitions of causality can differ. For example, manipulation-neutrality can be the nature of causal relations; i.e., by changing a cause, you influence its effect, but such relations can be epistemically defined in terms of probability raising. The pursuit of the economists' understanding of causality employing the referentialist semantics is definitely limited to uncovering the latter type of understanding.

Furthermore, several research methods can help in uncovering relations depicted by more than one approach to causality. Alternatively, there is some disambiguity in the meaning of causality presupposed by each method. In some instances, despite assigning each method to a chapter devoted to one theory of causality, I will highlight other possible interpretations.

At a previous stage of my research, I have systematized and described research methods employed by economists to put forth causal conclusions employing systematic literature review of the research published by three top economic journals (*American Economic Review*, *Quarterly Journal of Economics*, and *Journal of Political Economy*) between 2005 and 2015 (Maziarz 2018). In the book, I exemplify the most common methods of causal inference with in-depth case studies chosen from the sample and deliver (brief) examples of those methods that are not popular today. Each case study proceeds as follows. First, I briefly review the context of the study that specifies empirical and theoretical research on the topic under consideration. Second, I analyze the research method employed by the authors of the article under consideration. Third, I study the types of causal relations that can be discovered by used methods. Fourth, I consider what the economic world would have to be like for the conclusions to be justified and reconstruct the presuppositions implicitly accepted by the use of each research method. Finally, I discuss types of policy interventions justified by evidence and their limitations. Here, I need to highlight that I do not want to repeat the known issues such as, for instance, the assumptions of estimation techniques, but consider the causal context of research explicitly.

1.3 The structure of the book

Apart from the front matter and introduction devoted to presenting the purpose of the book and discussing the method employed to studying causal inferences in economics, each chapter (Chapters 2–6) focuses on one of the main five philosophical approaches to causality. Each of them start by delivering a definition of causality representative for the views of the philosophical theories of causality belonging to the approach under consideration, and proceeds by reviewing in-depth studies of recent causal economic research. The case studies and reviews of philosophical debates allow for reconstructing views on causality presupposed by the use of a given research method and indicating how such evidence can be employed for policymaking.

The first part of each chapter serves the purpose of introducing the reader into each of the five main approaches to/definitions of causality and reviewing hitherto philosophy of economics discussions. I summarize the views of philosophers of causality intending to present a consistent definition (meaning) of causality, discuss the reception of each approach to causality among the philosophers of economics and economists, and summarize the voices present in the methodological literature. The second part of each chapter focuses on discussing research methods that support causal claims understood in line with the approach under consideration. Each method of causal inference employed by contemporary mainstream economists is exemplified with research published recently in top economic journals. On average, I conduct three in-depth case studies in each chapter, but also give examples of other research methods. All examples of economic research are based on studies published by the three top economic journals (*American Economic Review*, *Journal of Political Economy*, and *Quarterly Journal of Economics*) from 2005–2015. The third part of each chapter focuses on discussing the policy-oriented implications of the case studies from previous sections. I consider the strengths and limitations of each method of causal inference in the context of using its results as evidence for policymaking. Based on the philosophical presuppositions underlying the research methods, the limitations of these methods, and the meaning of causal claims, I analyze what types of interventions are justified by the evidence resulting from particular methods of causal inference. Also, I discuss the misuses of causal evidence and reasons why even justified uses of causal evidence can result in unsuccessful interventions. Each chapter concludes with a summary.

Chapter 2 focuses on the regularity theories of causality and research methods grounded in this understanding of causality. In the first part, I briefly review the development of the regularity view on causality. David Hume, who defined causality in terms of constant event conjunctions, is the father of this reductionist position. Due to criticism, John Stuart Mill added the requirement of necessary connections between conjoined events. In the first half of the twentieth century, logical positivists revived the reductionist view. The regularity approach was further developed by defining causes as difference-making factors, or INUS conditions. I also discuss philosophical views of econometricians from the Cowles Commission. In Section 2.2, I focus on case studies of research methods employed by economists. The main research methods aimed at uncovering constant event regularities are structural equation modeling (SEM), theory-driven econometrics, and two cliometric studies of economic history (an analysis employing narrative records and statistical analysis). In Section 2.3, I discuss how evidence delivered by the research methods can be employed to policymaking. I differentiate between theory-driven and data-driven regularities, and discuss why such evidence does not warrant causal claims to be invariant under intervention and its heavy reliance on theory. I summarize the chapter by differentiating between the two types of interventions and argue that only those policy actions that do not rely on 'translating' causal claims into the manipulationist view are vindicated.

Chapter 3 discusses the probabilistic approach to causality that reduces causal relations to changes in conditional probability. In the theoretical part, I summarize the criticism of the regularity views and discuss the developments of Wiener, Suppes, Good, and Cartwright. I also review philosophical issues in the debate on econometrics: the views of Keynes, Hicks, and Granger's development of Wiener's definition of causality as a change in predictability that gave birth to the project of atheoretical econometrics. The discussion of case studies covers an exercise in vector-autoregressive modeling as a conventional example of atheoretical econometrics and a cross-sectional model. The latter example shows that also regressions usually interpreted as pieces of theory-driven econometrics can be practiced in an atheoretical way; i.e., without establishing causal structure on aprioristic grounds. In Section 3.3, I focus on the use of atheoretical econometric studies as evidence for policymaking. My main concern is that such results are susceptible to the common-cause fallacy, or spurious correlations. To offer a practical solution to this problem, I distinguish between policy actions that do and do not break the causal structure and exemplify these two types of interventions with examples.

Chapter 4 focuses on the counterfactual theories of causality. In the first section, I review the theories put forward by Lewis, Mackie, and other counterfactual formulations of the necessary condition.[3] Furthermore, I consider the views of philosophers on the use of counterfactuals in research practice (voiced by Reiss, Cartwright, and Shadish, Cook, and Campbell) and philosophical discussion on the methods of case-study analysis. The second part focuses on the methods of causal inference aimed at justifying counterfactual claims: drawing counterfactual claims from a previously established calibrated model and a case study. In Section 4.3, I comment on the distinction between Galilean and manipulationist counterfactuals. The latter type of counterfactual claims does not warrant that an intervention on a cause will result in a change in its effect. This undermines some (but not all) types of interventions. I consider how policymakers can deal with this problem and conclude by discussing the vices and virtues of the counterfactual evidence.

Chapter 5 discusses the mechanistic approach to causality. As usual, in the first section, I focus on discussing mechanistic theories of causality with the view to distinguish between ontic and epistemic views. I also review various definitions of 'mechanism' and consider the concept of economic mechanism put forth by Marchionni. I also cover the mainstream philosophy of economics discussions regarding theoretical modeling that interprets theoretical models as models of mechanisms. Three methods of causal inference aimed at uncovering theoretical mechanisms are exemplified with in-depth case studies: inferring causal claims on a ground of a nonempirical theoretical model, calibrated theoretical model, and its special case, DSGE (dynamic stochastic general equilibrium) model. In Section 5.3, I discuss two main issues connected to mechanistic evidence: (1) the problem of empty mechanisms, and (2) the in-principle unpredictability of the results of interventions grounded in such knowledge. To end optimistically, I consider the type of interventions that is considered less often; i.e., designing economic mechanisms and deliver examples of successes and misuses of mechanism-based interventions.

Chapter 6 focuses on the last of the 'big five' approaches to causality. The first section starts by reviewing the manipulationist theories of causality and analyzing why they are usually considered as most adequate by philosophers interested in policymaking-oriented sciences. I consider the manipulationist reading of econometric modeling and discuss philosophical views on experimental and quasi-experimental research designs. The second part of the chapter focuses on the growing number of studies employing various experimental approaches to causal inference. I discuss an instrumental variable (IV) and natural experiment as quasi-experimental research designs and laboratory market and randomized field experiments. The third, policy-oriented, section focuses on the use of experimental and quasi-experimental evidence for policymaking. I differentiate between two types of extrapolation, and argue that the extrapolator's circle can be solved by conducting contextualized experiments that address policy questions.

Chapter 7, which plays the role of a summary, argues in favor of a thesis that economists as a group are conceptual pluralists: they use various methods of causal inference that allow for formulating causal conclusions understood in line with different notions of this relationship. These different types of evidence support causal claims based on different notions of causality, and therefore having different policy implications. I put forward the view that not each kind of causal evidence can justify interventions understood in line with the manipulationist definition of causality. In other words, I argue that translating causal claims into the manipulationist notion is not justified, and may lead to unexpected policy outcomes or failed interventions. Finally, I discuss unsolved problems and indicate the areas of further research.

Notes

1 The majority of these understandings of causality have been developed with the view to informing the discussions focusing on the natural sciences (mostly physics). Therefore, the five standard approaches are not always adequate to depict causality in the social sciences. Specifically, understanding causality in terms of physical processes (e.g., a transmission of energy from a cause to its effect) is inadequate with the picture of social ontology delivered by social sciences and economics. One of the formulations of the causal-process theories is the mechanistic approach that, roughly speaking, exchanges the demand of energy transfer with mechanisms underlying causal relations.

2 Here, I use 'model' in the widest sense that covers research methods ranging from both theoretical and empirical models to various forms of experimental and quasi-experimental designs.

3 I should note here that the counterfactual formulation of manipulationist theories (e.g., supported by Woodward) is discussed in Chapter 6.

References

Aristotle (1999). *Physics*, edited by D. Bostock. Oxford: Oxford University Press.

Bacon, F. (1852). *An Explanatory Version of Lord Bacon's Novum Organum*. London: Orphan School Press.

Beebee, H., Hitchcock, Ch., & Menzies, P. (2012). Introduction. In: Beebee, H., Hitchcock, Ch. & Menzies, P. (eds.) *The Oxford Handbook of Causation* (pp. 1–18). Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199279739.001.0001

Bianchi, A. (2015). Introduction: Open problems on reference. In: Bianchi, A. (ed.) *On Reference* (pp. 1–18). Oxford: Oxford University Press.

Carroll, J. (2009). Anti-reductionism. In: Beebee, H. et al. (eds.) *The Oxford Handbook of Causation* (pp. 279–298). Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199279739.001.0001

Claveau, F., & Mireles-Flores, L. (2014). On the meaning of causal generalisations in policy-oriented economic research. *International Studies in the Philosophy of Science*, *28*(4), 397–416. DOI: 10.1080/02698595.2014.979669

Claveau, F., & Mireles-Flores, L. (2016). Causal generalisations in policy-oriented economic research: An inferentialist analysis. *International Studies in the Philosophy of Science*, *30*(4), 383–398. DOI: 10.1080/02698595.2017.1331976

Converty, A. (2006). *Hume's Theory of Causation: Quasi-Realist Interpretation*. London: Continuum International Publishing Group.

Frege, G. (1948 [1892]). Sense and reference. *The Philosophical Review*, *57*(3), 209–230. DOI: 10.2307/2181485

Gauker, Ch. (2015). The illusion of semantic reference. In: Bianchi, A. (ed.) *On Reference* (pp. 21–39). Oxford: Oxford University Press.

Henschen, T. (2018). The in-principle inconclusiveness of causal evidence in macroeconomics. *European Journal for Philosophy of Science*, *8*(3), 709–733. DOI: 10.1007/s13194-018-0207-7

Hoover, K. (2001). *Causality in Macroeconomics*. Cambridge: Cambridge University Press.

Hoover, K. D. (2004). Lost causes. *Journal of the History of Economic Thought, 26*(2), 149–164. DOI: 10.1080/1042771042000219000

Lakatos, I. (1980). *Mathematics, Science and Epistemology: Volume 2, Philosophical Papers*. Cambridge: Cambridge University Press.

Lossee, J. (2012). *Theories of Causality: From Antiquity to the Present*. Livingston: Transaction Publishers.

Maziarz, M. (2018). Causal inferences in the contemporary economics. *Mendeley Data*. Retrieved from: http://dx.doi.org/10.17632/v7dhjnd8xg.2. Access: 16th October 2018.

Mill, J. S. (1893). A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation. *Harper &Brothers*. Retrieved from: ftp://skipjack.soc.lib.md.us/pub/gutenberg/2/7/9/4/27942/27942-pdf

Reiss, J. (2012). Causation in the sciences: An inferentialist account. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(4), 769–777. DOI: 10.1016/j.shpsc.2012.05.005

Reiss, J. (2015). *Causation, Evidence, and Inference*. London: Routledge.

Serene, E. F. (1979). Robert Grosseteste on induction and demonstrative science. *Synthese*, *40*(1), 97–115.

Tooley, M. (1990). Causation: Reductionism versus realism. *Philosophy and Phenomenological Research*, *50*, 215–236. DOI: 10.2307/2108040

Woodward, J. (2015). Normative theory and descriptive psychology in understanding causal reasoning: The role of interventions and invariance. In: Gonzalez, W. (ed.) *Philosophy of Psychology: The Conception of James Woodward* (pp. 71–104). Berlin: De Gruyter.

2 Regularities

The 'regularity approach to causality' label dates back to David Hume, an eighteenth-century philosopher who started the modern debate on causality, but the label also covers developments voiced by philosophers of causality in the second half of the twentieth century. Generally speaking, the philosophers ascribing to this understanding of causality agree that variously understood regularities or constant conjunctions of events are signs of or are themselves causal relations. Four different types of regularity theories can be distinguished. First, causality can be understood as two events being constantly conjoined; i.e., appearing at the same time and space. Second, causal relations can be defined as empirical regularities instantiating 'necessary connection,' i.e., regularities produced by a law of nature. Third, causes are, in some cases, defined as difference-making factors, or the INUS conditions. Fourth, according to the logical-positivist reductionism, there is nothing beyond appearances in laws of nature (the regularity view of laws, or RVL). This chapter dives into these four groups of philosophical views belonging to the 'regularity approach' (Section 2.1), studies how economists uncover empirical regularities (Section 2.2), and discusses the use of such evidence for policymaking (Section 2.3).

2.1 Reducing causality to constant conjunctions

The regularity approach to causality seems to be out of fashion today, at least considering the topics present in the current philosophical debates. However, defining causality in terms of constant conjunctions had been popular at the beginning of the twentieth century when logical positivism was the mainstream stance in the philosophy of science and econometrics had been established. Even though philosophers today, generally speaking, reject the regularity view, this approach to causality still shapes the quantitative branch of economics. The project aiming at reducing causality to constant conjunctions dates back to David Hume. The philosopher who started the modern debate on causality defined causal relations in terms of regularity: "[w]e may define a cause to be 'An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedence and contiguity to those objects, that resemble the latter'" (Hume 1956, p. 170). This passage is the cornerstone of the regularity account of causality.

According to the Humean view, a cause of an event always precedes its effect in time and is located in the same place. The relata of the cause–effect constant conjunctions are events (such as a ball striking another ball) (Stroud 2000, p. 28). The definition of the constant-conjunction approach to causality can be formalized as follows: *c* causes *e* if and only if

> i. *c* is spatiotemporally contiguous to *e*; ii. *e* succeeds *c* in time; and iii. all events of type *C* (i.e., events that are like *c*) are regularly followed by (or are constantly conjoined with) events of type *E* (i.e., events like *e*).
>
> (Psillos 2009, p. 131)

The definition delivered by Stathis Psillos can be interpreted as a reductionist view indicating that causality is nothing beyond the spatiotemporally conjoined events. However, in *An Enquiry Concerning Human Understanding*, Hume (1992, p. 51) supplemented the definition just quoted with the passage that can be read in line with the counterfactual approach (cf. Chapter 4): "[o]r in other words, where, if the first object had not been, the second never had existed." Hume connected them with the phrase "in other words" that suggests their synonymity, even though the two definitions are distinct.

What seems to be an inconsistency at first glance results from the distinction between ontic and epistemic definitions of causality. While the former aims to grasp what causality really is (or, to put it differently, what is the nature of the phenomenon), the latter limits its scope to reducing causality to such relations that can be observed. As Psillos (2009, p. 133) put it,

> we should distinguish the epistemic question of how we come to know the presence of a regularity, given that our evidence for it always has to do with past and present instances of it, from the metaphysical question of what kind of entity a regularity is.

Interpreting these two passages as distinct voices regarding the ontology and epistemology of causality saves Hume from being committed to accepting the inconsistency. Hume's other writings suggest that he was a realist about causality (Converty 2006; Wright 1973, 2000) and therefore he is likely to accept a pluralist stance (Loptson 1998). The counterfactual definition is an ontological view: without causes, effects would not happen (see Chapter 4). Hume's epistemic position indicates his pessimism about the ability of human senses to observe causation: we can only observe that one event *c* always follows *e* while the purportedly causal connection between the two events remains unknown.

2.1.1 From constant conjunctions to the regularity view of laws

Whatever Hume had in mind while putting forward the two definitions, the 'Humean view' is now identified with a reductionist approach according to which "there is no extra element in causation which is of a fully distinct kind, like a necessary connection or a productive relation . . . that would explain or ground or underpin the regular association" (Psillos 2009, p. 132). The opinions among philosophers of causation accepting the regularity account as an epistemic stance differ regarding the realism of causal relations. 'The Humean view,' i.e., reducing causal relations to constant conjunctions of events, is problematic because it is insufficient to offer a criterion for distinguishing between accidental and causal regularities. In other words, the opponents of the epistemic definition of causality put forward by Hume argued that it is only partially true: causal relations do indeed produce constant conjunctions of events, but there is something more beyond empirical regularities that differentiates causal and noncausal relations. Among others, Immanuel Kant criticized Hume for not explaining why regularities exist and argued that "every event is determined by a cause according to constant laws" (Guyer 1992, p. 219). A similar viewpoint was voiced by Thomas Reid (1815, p. 282). He argued that if causality is identified with observing that one event follows another, then "[i]t follows from this definition of a cause, that night is the cause of day, and day the cause of night" (cf. Gallie 2013). Such arguments also appear in the contemporary debates as voices against the regularity theories. For instance, Strawson (2014) argued that every empirical regularity seeks an explanation.

In response to such criticism, John Stuart Mill (2017) argued that a regularity indicates a causal relation only if a law of nature produces the regularity under consideration. The nineteenth-century philosopher and economist added the demand for necessary connection between cause and effect, intending to distinguish between accidental and causal regularities. For Mill, only those regularities that are produced by laws of nature are causal. Logical positivist philosophers popularized Mill's view by employing its modification to their account of scientific laws and explanation. Although one of the central tenets of the doctrine was to differentiate between science and nonscience (metaphysics), discussing explanation required a concept of causality, at least considering the contemporary state of the debate (Salmon 1989, p. 172).

In the beginning, neopositivists identified laws with regularities what is known as the regularity view of laws, or RVL for short. Opposing metaphysics, logical positivists accepted an ontologically reductionist definition of scientific laws. For instance, Rudolf Carnap (1966, p. 3), understood a scientific law as "a . . . regularity . . . observed at all times and all places, without exception." Neopositivists rejected the Millian 'necessary connection' as a circular notion: since laws are constant and exceptionless regularities, defining regularities cannot require being produced by laws. Scientific laws are crucial for the logical positivist account of explanation. Carnap (1966, p. 8) depicted explanation as the process of delivering a scientific law and conditions in which the phenomenon under consideration occurs. Carl Hempel (1965, pp. 335–338) developed Carnap's views by putting forth the deductive-nomological (D-N) and inductive-statistical (I-S) (pp. 380–384) accounts of scientific explanation. The use of statistical regularities in the latter model of explanation gives a hint that Hempel accepted not only strict, deterministic regularities, but also probabilistic relations as causal. Even though Hempel's account of explanation can be employed to explaining in noncausal terms (e.g., being a raven explains blackness of the bird behind the window), the Berlin-Circle philosopher highlighted that his model is aimed at serving for constructing causal explanations if "an appropriate law or set of laws holds by virtue of which [*c*] causes [*e*]" (Hempel 1965, p. 350).

Even though defining laws in terms of empirical regularities does not require the concept of causality, as Michael Stöltzner (2009) highlighted, "[c]ausation was a central theme for the movement of Logical Empiricism" beside the analytic-synthetic distinction and the confirmationist doctrine of theory appraisal. Physical developments of the nineteenth century inspired the reductionist view. For instance, Ernst Mach (1960), in *The Science of Mechanics*, voiced his belief that:

> [t]here is no cause nor effect in nature; nature has but an individual existence; nature simply is. Recurrences of like cases in which *A* is always connected with *B*, that is, like results under like circumstances, that is again, the essence of the connection of cause and effect, exist but in the facts.
>
> (p. 580)

Mach's voice can definitely be interpreted as a stance reducing causality to constant regularities in regard to epistemic view on cause and effect. However, it can also be read as a non-reductive view in the ontology of causality: existence 'in the facts' indicates that facts somehow supervene on causal reality (cf. Becher 1905).

Another member of the neopositivist movement, Philip Frank (1961, p. 63) explicitly referred to the Humean, reductionist viewpoint and defined causality as a situation where "a state of the universe *A* is once followed by the state *B*, then whenever *A* occurs *B* will follow it." Observing the following two implications of this definition is useful for the reconstruction of philosophical views presupposed by economists. First, such a definition is committed to a deterministic worldview: the same conditions will always produce the same results. Second, the definition is unverifiable: to confirm a claim that *A* causes *B*, one would have to observe that a state of the world *A* repeats itself, and such repetition is extremely unlikely. Therefore, Frank's proposition can hardly be refuted. In his later book, Frank (2004, p. 280) distinguished the Humean and Kantian definitions differentiating between them as follows: the former is based on the recurrence of a state of the world and the latter on scientific laws. Morris Schlick (1936) voiced similar views, but, crucially for economics, argued that statistical laws result from the lack of knowledge instead of inherent indeterminism.

Hans Reichenbach (1971) focused on developing methods of causal inference given the constant-regularity view is accepted at least as an epistemic stance. According to Reichenbach's (1971, p. 167 et seq.) views grounded in physics, causal relations are characterized by the spatiotemporal proximity and order: causes act by contact and precede effects. Such relations can be inferred using statistical analysis as follows. First, a law $F_{apriori}(c_1; c_2; \ldots; c_n)$ is presupposed. Second, similarly to the process of Bayesian econometric modeling (cf. Gelman and Shalizi 2012), new observations are employed in order to infer new, statistically significant causes. Finally, either the process ends and an empirical law $F_{empirical}(c_1; c_2; \ldots; c_m)$ is obtained, or the inductive procedure is a never-ending process when new observations infinitely allow for obtaining new parameters $(c_{m+1}; c_{m+2}; \ldots)$ what indicates that the law is a result of spurious correlation (Reichenbach 1971). Furthermore, Reichenbach (2012a, 2012b) received the attention of philosophers for establishing a cornerstone of the probabilistic approach to causality (cf. Chapter 3).

In summary, the Humean stance that reduces causal relations to constant conjunctions of events was reformulated by the logical-positivist philosophers who strived to deliver a reductionist view on laws of nature. As Stathis Psillos (2009, p. 141) put it, the regularity view of laws denies that "laws, as they are in the world, are anything over and above stable patterns of events . . . to call a sequence of events *c* and *e* causal is to say that this sequence is a part of (instantiates) a regularity." Today, this stance is labeled as the Ramsey-Lewis view (Beebee 2000, p. 571): *c* causes *e* if and only if the two events instantiate a law-like regularity. To restrain from using the concept of cause, laws are taken to be those generalizations that are the most economical and correct axiomatization of observations representing some phenomena.

2.1.2 Further developments

Arthur Pap (1952) observed that using the constant conjunction definition requires considering the similarity of events. Otherwise, no causal claims can be established since that would require a state of the world to repeat itself. He argued that "[a] difficulty . . . which has received insufficient attention in the heat of the debate about the meaning of 'necessary connection' is, just how the classes of 'constantly conjoined' events, relevant to a given singular causal judgment, are to be selected" (p. 657). Pap put forth the solution according to which the classes are defined subjectively, having a purpose of research in mind. However, too-broad or too-narrow definitions of classes lead to observing regularities full of exceptions or exceptionless and trivial regularities. In order to solve this dilemma, Curt Ducasse (1951) offered a singularist theory of causality. According to his approach, the cause is a difference-making factor in the direct spatiotemporal proximity of an effect. The American philosopher argued that *c* is a cause of *e* if and only if it is the only difference in the spatiotemporal surrounding of *e*. Both Ducasse's definition and the method of difference coined by John Stuart Mill define causes as necessary conditions.

In addition to extending the Humean view by adding the condition that regularities interpreted causally should instantiate a law of nature, Mill also worked on methods of causal inference. Among several different concepts, the method of agreement and difference are most notable for its use even today: comparative case studies (described in Section 2.2.3) are, roughly speaking, a method looking for the factor causing a difference between cases. The 'Method of Agreement' states that if phenomena *p* is caused by several groups of factors $(F_1, F_2, \ldots, F_n)$ and one of them (F_c) can be observed in each instance, then F_c is a sufficient cause of p. On the contrary, the 'Method of Difference' indicates that the comparison of two sets of circumstances (one that produced phenomena p and the others that did not) can uncover that a cause of p is a circumstance that was observed in the former but not in the latter case. In other words, the method of difference aims at uncovering necessary conditions. According to Stathis Psillos' (2009) interpretation, Mill believed that causes and antecedent conditions are indistinguishable and defined causes as sufficient conditions. This indicates his instrumentalist leaning.

A century later, John Leslie Mackie (1974) developed the regularity approach to causality by distinguishing between necessary and sufficient conditions and required from causes to be (at least) INUS condition. Mackie (1974) believed that the advantage of reducing causal relations to constant event conjunctions is that it "involves no mysteries" (p. 60). Sufficient factors are a group of causes that together can deterministically produce an event. However, each one, taken separately, would not produce the effect. Necessary factors are causes without which an effect would not be produced. Considering which group of factors should be labeled 'cause,' Mackie (1965) answered that causes are at least "insufficient but necessary parts of a condition which is itself unnecessary but sufficient for result" (p. 245). In other words, causes are (1) insufficient (they cannot produce an effect without the existence of other factors), (2) non-redundant (the effect would not be produced without the cause, *ceteris paribus*), (3) unnecessary (the effect can be produced without the cause, and (4) sufficient (given a context, they produce the effect). To grasp the concept of the INUS condition, think about a company that produces good *G* and sells it at an efficient market (is a price taker). In such a situation, an improvement in technology, lowering production costs, is the INUS condition of higher profit. Such a change is an insufficient cause because, without other technologies, factories, and organization structures, it would not affect profit. It is non-redundant: in a given context, higher profitability would not occur without technological change. However, the condition is also unnecessary: higher profits can be recorded without the technological change when, for instance, the market price changes. Finally, technological change is a sufficient cause of higher profitability: other things being constant, such a change brings about the effect. The probabilistic theories of causality (cf. Section 3.1) further liberalize the requirements and state that even those factors that are not INUS conditions, but only raise the probability that an effect occurs deserve the causal label (Suppes 1970, p. 34).

2.1.3 Criticism and rejection of the regularity view

The probabilistic approach replaced the regularity view on causality in the second half of the twentieth century. On the one hand, developments in quantum physics and the social sciences required a philosophical interpretation of probabilistic laws and made scientists demand a theory of causality adequate to their results. On the other, philosophers encountered a few problems with reducing causal relations to constant conjunctions, law-like regularities, or employing

INUS conditions. For example, the use of the constant-conjunction definition to causal inference requires operationalizing classes of events. To say that *c* and *e* are constantly conjoined, one needs to have a description of events under consideration. If the classes of events are defined too strictly, then events never recur while too broad definitions lead to results that lack informativeness.

Christopher Hitchcock (2010) listed three central problems. First, regularities in the real world are not perfect. The success of the Copenhagen interpretation of quantum mechanics and the rise of social sciences proved that many causal relations are described with statistical laws instead of exceptionless regularities. Second, the regularity account fails at distinguishing between difference-making and unnecessary conditions (event *A* can always be followed by event *B* even if *A* does not 'cause' *B* in the sense that by changing *A*, one can modify *B*), which makes it irrelevant to causal inferences. Third, the asymmetry of the causal relations is left without delivering explanations or more in-depth analyses beyond presupposing that causes precede effects in time. The consensus among philosophers of causality states that the regularity view has been rejected. However, I argue ahead that this view still influences the research practice in economics.

2.1.4 The regularity approach in the philosophy of economics

Historically, the time when logical positivism was the mainstream stance in philosophy of science correlated to the beginning of modern econometric techniques (cf. Keuzenkamp 2000, pp. 213 et seq.). However, not only the temporal relation binds specific research methods with the regularity theories. Some economists and philosophers of economics took part in the debate on the regularity approach to causality. Two notable examples of influencing the philosophy of causality literature (apart from John Stuart Mill's extension of the Humean definition of causality discussed previously) are Sheila Dow's (2002) voice in the discussion of the seemingly contrary definitions delivered by Hume and Daniel Hausman's (1998) attempt at establishing a new way of differentiating between causes and effects on the grounds of observational data. Others focused mainly on methodological issues connected to the project of econometrics rather than discussing causation in the abstract. Specifically, philosophically minded econometricians and philosophers of economics interested in the empirical branch of econometrics have long debated evidence determining the direction of causal inferences, the interpretation of error terms, the (in-)deterministic nature of economic regularities, what produces regularities in the economy, and other issues.

Sheila Dow (2002) reviewed the historical debate of various interpretations of the Humean thought, indicating that they differ due to being context-dependent and putting forth the argument that an appropriate reading should consider the presuppositions of the epoch and time when a viewpoint was coined. Using this method and considering the ideas of the Scottish Enlightenment, Dow (2002) interpreted Hume "as a rational skeptic, the inspiration for logical positivism, and a realist" (p. 399). In her later work, Dow (2009) analyzed Hume's influence on the modern economics listing three areas: the quantity theory of money, the influence of trade on development, and what is of particular interest for this book, his promotion of empiricism that supposedly shaped the project of econometrics. According to her argument, Hume's reductionist approach is the cornerstone of econometric research aiming at analyzing causal relations by means of studying empirical regularities in the observational data. Tony Lawson (1997), who promoted the critical-realist philosophy of science and opposed Hume's reductionism, agreed with Dow's interpretation. Another notable example is Hausman's (1998) attempt at solving the problem of inferring the direction of causal relations among contemporaneous events and variables (i.e., deciding whether $A \rightarrow B$ or $B \rightarrow A$ is true) by assuming that every event has at least two causes (cf. Psillos 2009). Hausman argued that the effects determined by common causes are correlated, but causes producing one event are not what makes causal inferences possible.

Econometrics: methodological discussions

The Cowles Commission established the causal interpretation of econometric models (particularly the structural models) in the first half of the twentieth century (cf. Morgan 1992; Pearl 2012). The discussion of whether structural-equation modeling (SEM) can be interpreted causally is far from being settled. Some philosophers argue that this approach is based on econometric regression, which, similarly to correlation, is a symmetrical relation and therefore heavily relies on theory to establish the asymmetry between cause and effect (e.g., Freedman 1987; Dawid 2010). Daniel Hausman (1983) also disagreed with the widespread interpretation according to which structural equation models can be interpreted causally. According to his view, the fact that some variables are located on the right-hand side of equations is not sufficient for imposing the causal interpretation on these models without additional evidence.

On the contrary, Cartwright (1989, pp. 149–150) argued that economics is an explicitly causal science from its beginning and therefore economic models should be interpreted as depicting relations between causes (located on the right from the equal sign) and effects. Also, economists (e.g., Adda et al. 2014; Artuç et al. 2010; Carneiro and Lee 2011) interpret structural models causally. This opinion is also shared by some philosophers (e.g., Woodward 2016; Pearl 2012) who accept the causal interpretation but differ regarding the meaning of causal claims.

The question of what is the meaning of causality presupposed by econometricians[1] has received considerable attention. Julian Reiss' (2009) opinion that econometric methods are grounded in the probabilistic accounts seems to be oversimplified unless econometrics is not narrowed exclusively to observational studies conducted in an atheoretical way (more on this issue ahead). Kevin Hoover (2008) distinguished between Wold's process-analysis, Granger-causality tests, and VAR (vector-autoregression) analysis that are grounded in the demand of time-precedence emphasized by Hume and non-reductive methods presupposing the manipulationist stance of Simon (1977 [1957]) and himself (Hoover 2001) (the so-called Hoover test). Also James Woodward (2016) accepts this position while acknowledging that "[m]y own answer is interventionist: an equation . . . correctly describes a causal relationship if and only if for some range of values of *X*, interventions that change the value of *X* result in changes in *Y* in accord with . . . [it]" (p. 16). Recently, Tobias Henschen (2018) analyzed cases of macroeconometric research and argued that macroeconomists understand causality in line with the manipulationist account. I and Robert Mróz argued that Henschen's definition is too narrow and adequate only to a sample of macroeconometric models, while other models represent relations deemed causal only if one accepts different definitions, and therefore causal pluralism seems to be adequate to macroeconomics (Maziarz and Mróz 2019). On the contrary, Alessio Moneta (2005) concluded that econometricians modeling macroeconomic phenomena are antirealists regarding causality, and reduce this concept to empirical regularities.

In my view developed in the following chapters, different econometric methods help discover causal relations understood in line with the regularity, probabilistic, and manipulationist definitions of causality. Whether econometric techniques aim at discovering constant regularities instantiating necessary connections, probabilistically understood causal relations, or intervention-invariant connections (changes caused by interventions) depends on utilizing additional knowledge.[2] Generally speaking, three different concepts of causality are implicitly presupposed by the project of econometrics: cliometric studies and theory-inspired econometric models presuppose a version of the regularity view; atheoretical modeling delivers knowledge on probabilistic dependencies; and quantitative analyses of experiments and quasi-experimental studies (e.g., natural experiments) can uncover manipulationist-neutral relations (cf. Figure 2.1).

There are no explicit causal relations in datasets. Interpreting correlational evidence in causal terms usually requires additional assumptions on what are the appearances of causal relations (e.g., time precedence) or/and further knowledge on phenomena producing data under consideration. The difficulties connected to drawing causal conclusions from quantitative analysis of observational data can be exemplified by estimating the empirical law of demand. Data do not determine if price causes quantity or if quantity causes price. In other words, the direction of the causal relation is not determined by data, and equations I and II are observationally equivalent:

$$\text{I:} \quad q_t = C_1 + \alpha_1 p_p + \varepsilon_1$$

$$\text{II:} \quad p_t = C_2 + \alpha_2 q_t + \varepsilon_2$$

Where:

q_t = quantity of a good sold at time t

p_t = an average price of a good sold at time t

ε = error term

α_n = estimated parameters

C_m = constants

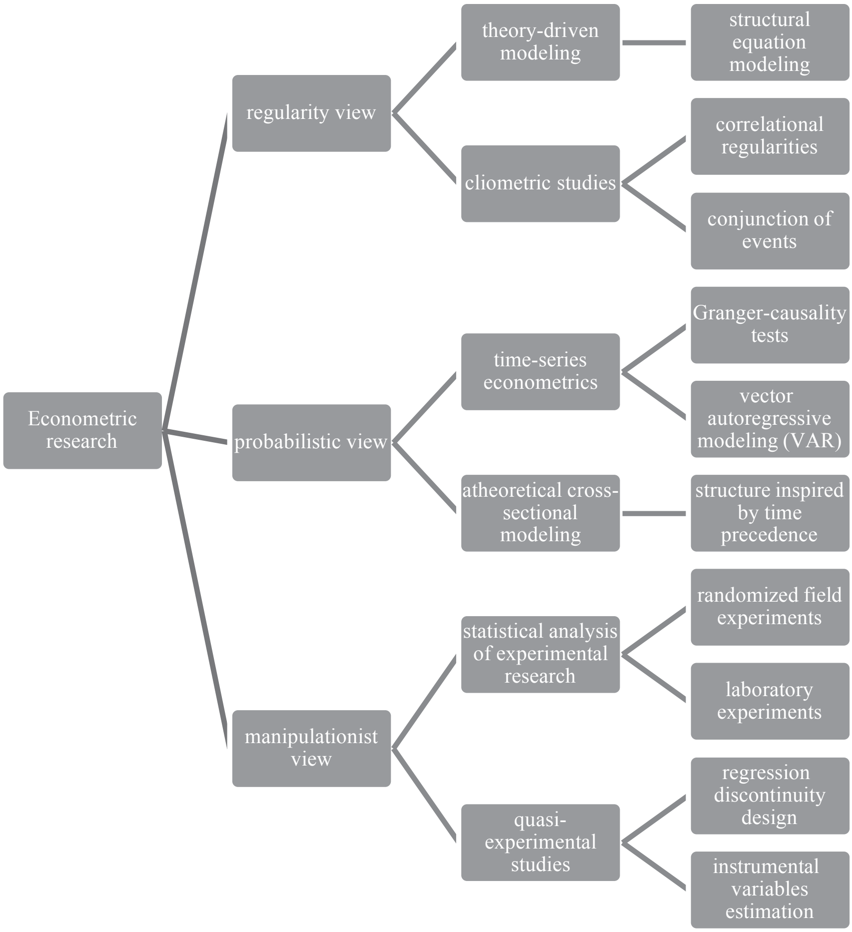

Figure 2.1 Econometrics and the meaning of causality. An overview

Without going beyond the observational data, the two equations are entirely equivalent. As I argue ahead, the type of additional knowledge (or lack thereof) used to identify causal structure determines the meaning of causal claims. Generally speaking, there are three different approaches to solving the problem of identifying the causal structure. First, econometricians can use aprioristic knowledge from economic theory. Second, they can inform the discussion by considering the actual interventions. Third, they can employ knowledge of time precedence or atheoretical knowledge on how data were produced. These three different approaches to causal inference differ in the presupposed meaning of causality.[3]

The chronologically first solution, known as the Cowles Commission approach, was developed in the 1940s in the context of structural equation modeling (cf. Section 2.2.1). The Cowles Commission approach employs aprioristic theories (models of phenomena, to use the contemporary dictionary) to determine the directions of causal relations. In other words, the causal structure is taken for granted as it is depicted by theoretical models. Only the parameters denoting the linear influences of each of the exogenous variables on the endogenous variable are estimated. In other words, causal relations depicted by econometric modeling conducted in line with the Cowles Commission approach instantiate 'necessary connection': the law-of-nature connection between cause and effect is depicted by economic theory.

Considering the case of an empirical law of demand, the influence of price on quantity of goods sold would be established on theoretical grounds, while the influence of a change in price on a change in a quantity known as the elasticity of demand was estimated. This interpretation of the structural equations, despite acknowledging the definition of causality as regularity, accepts that these regularities are generated by mechanisms or instantiate necessary connection in line with the adjustment to the Humean definition voiced by John Stuart Mill discussed previously. In other words, the aprioristic Cowles Commission methodology advises informing the choice between the two specifications of the empirical law of demand by considering the economic theory. The realist commitment is visible in Trygve Haavelmo's famous quote regarding car speed and position of the gas throttle (Haavelmo 1944, pp. 27–28). Moneta (2005, p. 440) indicated that:

> [i]n its declared objectives, the Cowles Commission methodology is non-reductionist and realist. There is something in reality at the macro-level that distinguishes autonomous (Haavelmo) or causal (Simon) relationships from empirical regularities. . . . [The Cowles Commission methodology assumes] the possibility of measuring causal properties of objective economic processes.

Despite criticizing this stance, Cartwright (1995) agreed to interpret regularity laws in causal terms and indicated that they are grounded in the reductionist empiricism inspired by David Hume. The regularity account of causality accepts laws that "can be either universal, in which case the law is deterministic, or they may be merely probabilistic" (p. 275). This proves crucial for the interpretation of econometric (structural-equation) models put forth by the Cowles Commission.[4] According to the standard Cowles Commission approach, the structural-equation modeling proceeds from theoretical analyses indicating the causes and effects of phenomena (exogenous and endogenous variables) to estimating coefficients of regressions.

The aprioristic approach to econometric modeling raises serious problems. On the one hand, relations between variables that are not depicted by economic theories cannot be modeled. On the other hand, it puts too much faith in economic theories that are, at least potentially, fallible. Herbert Simon (1977 [1957]) offered a solution based on employing knowledge about interventions to establish the direction of causal relations. Later, Hoover (2001) developed this approach and also opposed the a priori theorizing, cf. Chapter 6. Although Simon's opposition against the heavy reliance of econometric modeling on economic theory indicates his implicit support for the manipulationist approach to causal inference, his earlier views (Simon 1951) are in line with the Cowles Commission interpretation of econometrics.

As Simon (1977, p. 49) put it, "[i]n careful discussions of scientific methodology, particularly those carried on within a positivist or operationalist framework, it is now customary to avoid any use of the notion of causation and to speak instead of 'functional relations' and 'interdependence' among variables." The stance of the Commission is that the concept of causality is present in the philosophical debate since Aristotle and, despite criticism, is essential for establishing the asymmetry of (nonfunctional) relations. Opposing Hume, Simon did not found the time precedence to be essential but argued that causal relations should be identified using the asymmetry of cause and effect: "time sequence does, indeed, sometimes provide a basis for asymmetry between A and B, but . . . the asymmetry is the important thing, not the sequence" (p. 51). Therefore, for Simon (1977), it is the exogeneity and endogeneity of variables that is crucial for understanding the asymmetry between cause and effect. He exemplified this view with the case of relation between poor growing weather (W) and the price of wheat (P): the price depends on whether $P \leftarrow F(W)$, but the statement that "by changing the price of wheat we can affect the weather" (p. 52) $W \leftarrow F(P)$. is false. On these grounds, Simon (1977 [1957]) rejected the interpretation of econometric laws as functional dependencies and put forth the causal reading: exogenous variables placed on the right side of an equation are causes of the endogenous variables located on the left side of the equation.

Although the causal interpretation of the structural equations is the widely accepted mainstream stance, some thinkers disagree with it. For instance, Patrick Suppes (1970, p. 46) disagreed on the grounds that the causal structure implied by such models sometimes violate the assumption of time precedence that dates back to Hume. Additionally, the causal interpretation of structural equations is accused of not enlightening the real causal structure of phenomena. Christoper Sims (1980) criticized how the Cowles Commission econometrics was practiced for the heavy dependence on aprioristic assumptions, which indicated the direction of causal relations: "the style in which "identification" is achieved for these models . . . is inappropriate" (p. 1). He advised exchanging the structural models with data-driven approach methods such as vector autoregressive (VAR) modeling (cf. Section 3.2). Damien Fennel (2011, p. 372) criticized the causal interpretation for not justifying drawing manipulationist conclusions as follows:

> [d]espite its 'causal' label, as it stands it says nothing about the content of 'cause.' . . . Simon states that the exogenous variables should be taken to denote factors that are directly controllable by an 'experimenter' or 'nature,' and endogenous variables taken to denote factors that are indirectly controllable.

In other words, Fennel accused Simon's views of being based on aprioristic assumptions rather than proofs that the mechanisms modeled by the equations are, in fact, invariable under intervention. A glance at Simon's assumptions proves that Fennel's criticism is justified. As definition 3.7, (Simon 1977, p. 57) states:

> [l]et β designate the set of variables endogenous to a complete subset *B*, and let γ designate the set endogenous to a complete subset *C*. Then the variables of γ are directly causally dependent on the variables of β ($\beta \rightarrow \gamma$) if at least one member of β appears as an exogenous variable in *C*. We can say also that the subset of equations *B* has direct precedence over the subset *C*.

Simon (1977) highlighted himself that his method of the causal interpretation of the structural econometric models is of limited use regarding causal inferences when no theoretical grounds for a causal ordering is accessible since the conventionalist problem occurs. Usually, there is more than one set of equations that agrees with a set of observations. As Simon (ibid., p. 63) put it, "[a]n important objection to our definition of causal ordering . . . is . . . [that] the same set of observations could be represented by different structures with different causal orderings of the variables." To solve this problem, Simon (1977) employed the manipulationist meaning of causality (cf. Chapter 6).

The question if econometric models can be interpreted as causal models is not the only topic under consideration. The philosophical debates on econometrics also focus on what creates the regularities observed by means of econometrics if we can obtain knowledge exceeding appearances, and even whether empirical regularities exist in the social realm. On the one hand, the logical-positivist reductionist approach dominated the discussion in the past (cf. Cartwright 1989, p. 149). On the other hand, empirical regularities are read today as evidence for metaphysically rich concepts of causal relations such as the capacities approach (Cartwright 1995; Reiss 2013, p. 24; Earman et al. 2012). According to the capacity view on causality, causal relations supervene on entities' powers to behave in a particular way, but these capacities can stay inactive in certain circumstances (what resembles the behavior of social mechanisms, cf. Chapter 5). Such views, in their extreme form, lead to the conclusion that there are virtually any constant empirical regularities in economics. For instance, Łukasz Hardt (2017) argued that 'the dismal science' deals with reality devoid of empirical regularities excluding a few truistic laws such as Fisher's equation.[5] A similar position was voiced by Tony Lawson (1997). Contrary to Hardt's approach, Lawson started the argument from assuming a social-ontology position excluding the existence of constant regularities and argued based on this that the project of econometrics, which is aimed at discovering these purportedly absent regularities, is doomed to failure. According to critical realism, the openness of the social world makes it lack any empirical regularities unless experimental closure is artificially introduced.

The methodologists accepting regularity and probabilistic approaches to causality differ in regard to the interpretation of the error term ε. For Simon

(1957) and the Cowles Commission, ε assigns the aggregate influence of all the causal factors excluded from a model. This presumption gives a hint on the deterministic reading of causal relations held by the structural-equation modelers. According to this stance, if the out-of-model factors were excluded, then there would be deterministic relations observable in the economy. That is, the uncertainty introduced by the error terms is an effect of factors excluded from models instead of being produced by inherent uncertainty. On the contrary, the supporters of atheoretical methods presuppose inherent indeterminacy of causal relations and interpret them using the notion of statistical laws. Moreover, Cooley and LeRoy (1985) highlighted that the Cowles Commission approach to macroeconomic modeling underpinned the Keynesian empirical models. In contrary, the methods of quantitative causal inferences grounded in the probabilistic approach are connected to mainstream neoclassical macroeconomics. The Cowles Commission approach to econometrics states that if the factors included in the model could be experimentally isolated, then constant conjunction would be observable. On the contrary, the atheoretical econometrics implicitly employ the probabilistic definition (cf. Chapter 3) and therefore accepts the inherent indeterminacy of economics. However, this view is not new and can be traced back to Mill, who indicated that the operation of other factors makes economics fallible and inexact and forces the use of the *ceteris paribus* clause (Hoover 2008). The use of the CP-clause, according to Mill, is driven by our inability to access the mechanisms connecting cause with effect; i.e., "the origin and actual production of the phenomenon" (2017, p. 569), and the fact that causal antecedents do not always produce their effects; i.e., indeterminism of causal mechanism.

2.2 Establishing constant regularities

Economists infer causal claims employing research methods that are capable of discovering constant event conjunctions or law-like regularities present in the economy. This section discusses these methods and exemplifies them with case studies of contemporary economic research.

Three main approaches can be distinguished. First, theory-driven econometric research aims at uncovering empirical regularities instantiating Mill's necessary-connection requirement (or its contemporary counterpart; i.e., a realist view on scientific laws). I instantiate this approach with the case study of analysis conducted by Nicholas Bloom et al. (2012). Their method can be labeled as theory-driven econometrics, whereby the causal structure is identified using aprioristic modeling, and the strength of considered relation is estimated with econometric techniques from observational data. I also consider structural equation modeling that, despite being rarely used today, is of historical importance, and instantiate this method with the study of Erhan Artuç et al. (2010). Second, I analyze cliometric methods aimed at establishing the causal generalization in line with the constant-event-conjunctions view. Valerie Cerra and Sweta Saxena (2008) employ narrative records and quantitative techniques to analyze the relationship between economic crises and subsequent economic development. Third, I analyze a cliometric study presupposing the regularity view of laws. Finally, I discuss methods aimed at discovering difference-making factors and grounded in Mackie's INUS condition.[6] The following section addresses the question of how evidence delivered by these types of studies can be employed for policymaking.

2.2.1 Econometric research

Econometric methods were established at the time when logical positivism and the regularity approach to causality had been the mainstream philosophical positions. Therefore, older quantitative techniques are inspired by this philosophical stance. Precisely, the structural equation modeling and single equation regressions that import causal structure from theoretical investigations presuppose the meaning of causal relations as empirical regularities (correlations of variables) that instantiate 'necessary connection' or, in other words, are described by an economic, theoretical law.

Structural equation modeling

The Cowles Commission approach to causal inference is both theory- and data-driven. In detail, the causal structure is established on the grounds of economic theory, while coefficients denoting dependencies among variables are estimated from observational data. In other words, theoretical, aprioristic models inform the causal structure: what variables are causes and what variables are their effects. Econometric estimation delivers coefficients denoting the strength of causal relations between variables. For simplicity, the structural-equation modeling can be exemplified with the following model of self-contained structural equations:

$$\text{I} \quad a_{11}{\star}x_1 = a_{10}$$

$$\text{II} \quad a_{21}{\star}x_1 + a_{22}{\star}x_2 = a_{20}$$

$$\text{III} \quad a_{32}{\star}x_2 + a_{33}{\star}x_3 = a_{30}$$

Where:
a_{ij} = a parameter for *i*-th equation and *j*-th variable;
x_j = a *j*-th variable.

This system of equations implies that x_1 is an exogenous variable; on the contrary, x_2 and x_3 are endogenous variables (i.e., they are defined by values of exogenous variables). Considering the relation of determining between the equations, such a structure implies the following causal order: I → II → III: *x_1 is a direct cause of x_2, and x_2 of x_3* (Simon 1977, p. 58). The causal relation is defined in terms of determining a value of effect by its cause. Here, time precedence is not a key concept for establishing causal relations, however; it follows from the appropriate specification of the structure of equations. As Simon (1977) put it, "[t]here is no necessary connection between the asymmetry of this relation and asymmetry in time, although an analysis of the causal structure of dynamical systems in econometrics and physics will show that lagged relations can generally be interpreted as causal relations" (pp. 73–74). In a similar vein, Cartwright (1995) argued that the *ceteris paribus* laws (e.g., the law of demand) are grounded in the empiricist philosophy that dates back to Hume and the regularity account of laws and therefore can be read in causal terms. In fact, the law of demand could be reformulated into a two-equation structural model.

The structural-equation modeling *sensu stricte* seems to be an outdated research method: it is rarely used by contemporary economists to drawing causal conclusions (cf. Maziarz 2018). However, it is interesting to observe that research methods that are no longer in use among top economists are becoming popular in other disciplines. An excellent example of such methods is Granger-causality tests[7] that win the hearts of neuroscientists and biologists (Maziarz 2015). However, structural equation modeling is also used in natural and medical sciences. For instance, Kenneth Kendler, Charles Gardner, and Carol Prescott used this approach to quantitative analysis to study risk factors of depression. Considering that my book focuses on economics, I do not want to review their method in detail, but only cite their honest opinion on the limitations of the method they employed. The structural-equation modeling, or, in fact, the majority of curve-fitting econometric techniques, does not allow for inferring the asymmetry between cause and effect. The asymmetry (causal structure) needs to be established on theoretical grounds exceeding information included in the datasets or time precedence. Therefore, using such models in areas where either theory is underdeveloped, and non-informative or two competing approaches suggest alternative causal structures may lead to erroneous results. The problem rarely considered by economists has been noticed by a group of psychiatrists attempting to put forward an epistemological model of depression-inducing factors. Working on the quantitative analysis of determinants of depressive disorder, Kendler et al. (2006) acknowledge this fact explicitly when they state that "[s]ome of the intervariable relationships that we assumed take the form of $A \rightarrow B$ may be truly either $B \rightarrow A$ or $A \leftrightarrow B$" (p. 122).

Due to the limited popularity of structural equation modeling (at least in the contemporary mainstream economics), the discussion of SEM will be limited. However, it is useful to grasp the idea behind this approach because economists seem to use its informal version in their contemporary practice when they utilize theoretical knowledge to construct single-equation regression models (cf. 'theory-driven econometrics' discussed ahead). James Woodward (2005, p. 315) noticed the similarity of the two research designs which highlighted that both SEM and single-equation regressions inspired by theory represent causal structure 'imported' from theoretical considerations.

Erhan Artuç et al. (2010) employed the framework of structural equation modeling to address the question of what is the influence of trade shocks on workers' costs of moving between sectors. The econometricians used a dynamic equilibrium model of costly labor adjustment studied by Cameron et al. (2007). From the theoretical model, Artuç et al. (2010) deduce a reduced-form nonlinear regression and estimate it on a self-constructed dataset[8] covering statistics on workers' inter-sector mobility. While the general equilibrium model allows for hypothesizing that trade liberalization increases the demand for workers in sector profiting from trade shock, the econometric study delivers evidence for observing "sharp movement of wages in response to the liberalization" (p. 1042). Considering that the meaning of causality reconstructed by means of referentialist semantics in this study and the much more popular today approach of theory-driven econometric modeling is the same, motivated by the clarity of the argumentation, I focus on analyzing the less formalized study of Nicholas Bloom et al. (2012).

Nevertheless, I should highlight that, despite the limited popularity in contemporary economics (Maziarz 2018), structural-equation modeling is still widely debated in the methodology of economics. The debate is far from being settled. Recently, Stephen LeRoy (2018) differentiated between structural and reduced-form models and argued that only the former type could be interpreted causally because the latter models do not depict implementation-neutral (IN-) causality (cf. LeRoy 2018). The topic is also considered by philosophers interested in the methodology of natural sciences (e.g., Shipley 2016).

Theory-driven econometrics

Structural-equation models are deduced from theoretical models and therefore are highly formal and complicated mathematically. This approach to econometric modeling, despite being still widely discussed among the methodologists and philosophers, is rarely used by economists (Maziarz 2018). One of the reasons is that these models perform very poorly when used for predicting (Allen and Morzuch 2006) comparing to single-equation regressions and the atheoretical approach to econometrics (e.g., vector-autoregressive models) considered in the following Chapter 3. Another reason for its fall out of fashion is lowering the popularity of constructing general equilibrium theoretical models. For instance, Daniel Hamermesh (2013) delivered evidence for the 'empirical turn' in economics understood as a lowering the number of purely theoretical models and the shift of emphasis onto econometric and experimental studies. Possibly, contemporary economic theory, at least in some areas, does not have today the nature of large general-equilibrium structural models, but gathers various theoretical studies in an informal way (the phenomenon studied under the label of model pluralism; see, for example, Rodrik 2015). Many econometric studies no longer rely on the usual structural-equation methodology. Instead, econometricians either 'ground' their econometric model in a previously established theoretical model or construct such an aprioristic model on their own. The econometric model, contrary to the approach discussed previously, is not deduced from the theoretical model. Instead, one of the implications of the theoretical model (usually referred to as 'proposition') or previous empirical studies is mathematically transformed into information on the causal structure of phenomena, and econometric estimation is used to quantify the strength of causal relation.

The study of Bloom et al. (2012) can be labeled as the 'theory-driven' econometric modeling. On the grounds of constructed theoretical, aprioristic model and estimated econometric model, the authors conclude that "social capital as proxied by trust enhances aggregate productivity through affecting the internal organization of firms" (p. 1701). In other words, Bloom et al. (2012) seem to produce the following claims: (1) 'social capital' [trust, T] influences 'the internal organization of firms' (O), and (2) 'internal organization of firms' affects 'productivity' (P); what can symbolically be transcribed as $T \rightarrow O \rightarrow P$. The 'internal organization of firms' is a mediating variable. For clarity, I would like to focus on the previous causal claim, which was voiced more explicitly by the authors in the introduction where they admitted that the article presents "evidence that high social capital in an area increases decentralized decision making within firms" (p. 1663). In detail, the authors show that higher levels of trust are related to the higher decentralization of companies. This causal generalization refers to the results of the empirical model, while the theoretical analysis serves the purpose of showing that the empirical regularity is not spurious (instantiate a law of nature/necessary connection) in agreement with the distinction between causal and non-causal regularities voiced by John Stuart Mill or the contemporary, realistic view on scientific laws (cf. Section 2.1).

Before 2012, there were numerous studies analyzing the influence of culture on economic activity. The research on the influence of culture on economic activity dates back to the beginning of the twentieth century when Max Weber created his (1992 [1905]) opus magnum on the influence of reformationist culture on capitalism. For instance, the influence of various dimensions of societal culture on economic activity was studied econometrically (e.g., Tabellini 2010) and theoretically (e.g., Whiteley 2000). The authors quote the studies of Guiso et al. (2009) and Botazzi (2010) studies as examples of research on the influence of trust on economic activity. However, hitherto research focused on the macroeconomic (aggregate-level) phenomena. Bloom et al. (2012) explicitly admit that their analysis is "the first article looking at the role of trust on the organizational structure of *firms* across multiple countries, as opposed to country-level relationship" (p. 1667, emphasis in original). Previously, there were neither empirical nor theoretical studies aiming at analyzing the effects of cultural differentiation at the within-firm level.

The theoretical analysis aimed at establishing the relation between trust and company decentralization develops Garicano's (2000) model of within-firm knowledge acquisition. Bloom et al. (2012) study the decision of CEOs regarding delegating production decisions to plant managers. Rational managers are more likely to delegate decisions when the level of trust is high, because otherwise, plant managers are suspected by CEOs of being willing to act illegally. Considering the sign of the first derivative of function denoting the optimal level of decentralization, the authors deduce from the model that "[a]n increase in trust (λ rises) is associated with a higher degree of decentralization" (p. 1671). Considering the generality of assumptions, the model does not allow for deducing more than the sign and direction of the relation.

With a view to evaluating the strength of the dependency between trust and decentralization, Bloom et al. (2012) econometrically analyzed data gathered previously by interviewing managers of randomly chosen companies and quantifying the degree of decentralization. The measurements of trust were quantified on the grounds of World Values Survey data. While using widely recognized statistical constructs such as GDP also raises the same problem (which stays conventionally unnoticed), employing self-constructed measures lead to potential problems with the robustness of results. The vague concepts such as trust or the degree of company's decentralization can be quantified in many different ways and different operationalizations (measurements) of such concepts may lead to obtaining inconsistent results. Despite the widespread contemporary use of robustness analysis in econometric studies, they virtually never cover the alternative constructions of indices. As it will be shown in the last case study of this chapter, a change in how concepts are measured can change estimates or even reverse the relation.

Interestingly, Bloom et al. (2012) attempted to solve another type of possible fallacy: they repeated a fraction of interviews by choosing another manager from the same company and another team member to assess whether interviewed managers' and interviewees' biases systematically influence gathered data. The measures of decentralizations obtained in the two interviews were positively correlated ($\rho = 0.513$). Additionally, the authors validated the measure of trust by comparing it (p. 1691) with the values obtained by Geert Hofstede (2001). Their measure was positively correlated to the measure put forward by Hofstede.

The reliance of the econometric model on theory is explicitly admitted at the beginning of the empirical part of Bloom et al. (2012) paper where they state that "[o]ur theory predicts that greater trust of the CEO in the plant manager should lead to increases managerial delegation" (p. 1691). First, the economists used the most straightforward quantitative technique and calculated the correlation between the two variables. Trust and decentralization are positively correlated. The authors admit that the relation is 'highly significant' (p. 1692). Considering that they do not report statistical significance in this paragraph, they likely presuppose the distinction between economic and statistical significance (cf. Ziliak and McCloskey 2008) and report the former; i.e., they refer to a high degree of correlation instead of the high level of statistical significance.

Such a simple research design as calculating correlation is susceptible to the common-cause fallacy. In other words, other variables can cause both the dimension of culture and decentralization of companies. Furthermore, decentralization is also influenced by other factors. To control for the influence of those factors, the authors add several variables including firm size, skills, law, employment, and other characteristics of companies. Using the OLS[9] estimation, they arrive at the linear regressions (Table 1, p. 1692) for several subsamples. The econometric models can be presented symbolically as follows:

$$D_i = \alpha_t \star T_i + \sum_{i=1}^{n} \alpha_i \star X_i + \varepsilon$$

Where:

D_i = the decentralization z-score index for i-th company
α_t = the influence of trust on decentralization
T_i = the level of trust measured for the region where i-th company is located
$\alpha_i \star X$ = the influence of other determinants on decentralization
ε = error term denoting the summary influence of factors excluded from the regression

Running the regression for the whole sample allowed for estimating the parameters α denoting the influence of each variable. Considering that $\alpha_t = 1.231$, a 1 percentage point increase in the level of trust raises decentralization by 1.231 pp. The causal conclusion put forth by the authors refers to this result.[10] However, the dataset, taken separately, would not deliver sufficient evidence for the causal claim (even formulated in line with the regularity understanding). Let me consider the coefficient of correlation calculated by Bloom et al. (2012, p. 1692). The positive correlation between two variables (D; T) is not sufficient to indicate what is the direction of causal relation: both possibilities that $D \rightarrow T$ or $T \rightarrow D$ are equally plausible.[11] Therefore, the researchers establish causal structure on the grounds of theory (i.e., the aprioristic model established in the first part of their article). However, the theory is not only needed for informing econometric regression about the direction of causality. The theoretical model also delivers a theoretical justification for the empirical results. This justification allows for rejecting the purely correlational interpretation, and taking the results to be produced by a necessary connection, or – to put it differently – instantiate a law of nature (cf. Section 2.1). The research's inclination toward the regularity view on causality is clearly visible in the confirmationist views voiced in the introduction, where the authors take theoretical results as "predictions" for which they "find support from the hypotheses that trust increases decentralization" (p. 1664). Such a commitment seems to agree with the confirmationist philosophy of science.

I need to highlight that the Bloom et al. (2012) study is also liable to another reading. Namely, instead of interpreting it in line with the regularity approach to causality, the study could be read as an attempt at depicting mechanism connecting the dimension of culture and internal organization of companies employing theoretical modeling and confirming that this mechanism produces observable regularities. However, given that there are many mechanisms operating at the same time in the economy, which may screen each other off or multiply, one mechanism never produces an observable regularity. Hence, the knowledge of one mechanism is insufficient for putting forth testable hypotheses (cf. Chapter 5). Therefore, either economists' views on mechanisms diverge from the philosophical literature, or this reading is implausible. In Section 5.2, I show that economists constructing theoretical models implicitly accept this philosophy of economics mainstream view.

Furthermore, the authors do not regard their theoretical model as crucial part of the research or as a right description of mechanism while they admit that:

> [s]ince there are models other than our extension of Garicano (2000) that would predict a positive relationship between trust and decentralization we do not regard our empirical examination the final word on the correct theoretical model but as a useful framework for organizing our thinking.
> (p. 1675)

The 'correctness' of the theoretical model can be variously understood. If my view that by correctness, the authors mean 'being a model of true mechanism,' the passage indicates that the theoretical model presented in Section II is not purported to be a true account of mechanism, but only to give a hint on the connection between trust and company's structure decentralization.

With the view to substantiate the result, Bloom et al. (2012) estimated a few other models looking for the robustness of the result. In general, the robustness analysis falls beyond the scope of the book. Generally speaking, robustness analysis is conducted to show that the findings are not a result of data mining; i.e., small changes in estimation technique or variables do not influence the results significantly. Robustness checks make data-driven evidence more reliable. Therefore, robustness analysis supports the causal conclusion given the assumption that actual causal relations are not accidental but stable. What is peculiar, in this section, Bloom et al. (2012) used the instrumental variables (IV) estimation technique. Here, I need to distinguish between the following two uses of IV technique.

On the one hand, instrumental variables can be introduced to a model with a view to solving the (technical) problem of correlation between explanatory (right-hand-side) variables. On the other hand, the IV technique can be used as a method of designing a quasi-experiment (i.e., employing a semi-random sampling occurring independently from a researcher to measure the effects of an intervention under consideration). The latter use will be discussed in Chapter 6.

The causal claim cited at the beginning of the case study refers to the results of the theoretical model and econometric estimation. The causal structure (direction and existence of such relations) is established on the grounds of the theoretical model. The OLS econometric model estimates the strength of the influence of trust level on decentralization. If my reconstruction of the meaning of causality in the causal conclusion put forward by the authors is correct, then the study shows that there is an empirical association between the level of trust in society and firm organization so that companies operating within more trusting societies are more decentralized. Furthermore, this empirical regularity either instantiate necessary connection as understood by John Stuart Mill or the logical-positivist view on laws as economic descriptions of empirical observations. Bloom et al.'s (2012, p. 1664) mention of "other mechanisms" gives a hint on their acceptance of the realist reading of scientific laws (either in line with Mill's distinction between random and causal regularities, or a more recent philosophical view). Considering that laws of nature started to be interpreted realistically, as being produced by entities' features or capacities (Woodward 2016), the realist commitment of the authors is a good explanation for this passage.

2.2.2 Cliometrics

Cliometrics is a branch of economic history that employs various quantitative methods with a view to tracing the development of economic processes in time or generalize constant regularities. The methods vary from simple statistical techniques and basic econometric models to studies of co-existence (constant conjunctions) of events with the use of narrative records. Such studies are usually data-driven or, to put it differently, they limit the discussion of economic theory because either theory remains silent on a particular topic or different theoretical perspectives contradict each other, and therefore are uninformative.

In this section, I study two pieces of cliometric research employing different approaches to analyzing economic history. The first case is the study of Valerie Cerra and Sweta Saxena (2008) employing narrative records. The authors aimed at depicting the influence of economic recessions on long-term average economic growth. Carmen Reinhart and Kenneth Rogoff (2010) discovered an empirical regularity in the dataset covering average GDP growth and public debt.

The use of narrative records

Cliometricians use various techniques to identify events in economic history. While the usual approach is to study datasets and utilize knowledge on changing variables, a recent trend is to use various written (narrative) sources and transform them into useful statistics such as digital time series denoting the occurrences of events under study. For example, in the empirical debate on the 'expansionary fiscal contraction hypothesis,'[12] both approaches to identifying historical events appeared. Alberto Alesina and Silvia Ardagna (2010) defined fiscal contraction as a cut in government spending exceeding 1.5% of GDP and obtained the timing of major fiscal reforms on the grounds of statistical analysis of panel data covering government spending for each of the OECD countries. On the contrary, Daniel Leigh et al. (2010) decided to study budget plans and reports of international economic organizations. Christina Romer and David Romer (2010) employed a similar approach and analyzed various narrative sources such as presidential speeches and government reports to estimate the effects of tax policy on economic development.

A representative example of cliometric studies employing narrative records that draw causal conclusions is Valerie Cerra and Sweta Saxena's (2008) analysis of patterns in GDP growth after major economic slumps. "Growth Dynamics: The Myth of Economic Recovery" was published at the time when the most severe economic crisis since the Great Recession was developing. The authors addressed the following question: do recessions produce a long-lasting effect on the pace of economic development? In other words, they aimed at discovering whether GDP, after a recessionary shock, returns to its long-term trend.

As the authors indicated, at the time when the research was conducted, few studies focused on the question if countries experiencing major recessions recover from them and come back to their long-term growth rate. According to endogenous growth theories, long-term growth is determined by the pace at which total factor productivity grows (cf. Howitt 2010). Contradicting the received view, Cerra and Saxena (2008) admitted that their "paper documents that the large output loss associated with financial crises and some types of political crises are highly persistent," and their analysis delivers "some suggestive, although not definitive, evidence of causality" (p. 457). Concluding their research, the authors of the cliometric study highlighted that developing theoretical models of the phenomenon under consideration would be useful what gives a hint of their skepticism regarding direct application of their results to policymaking.

The causal conclusion refers to the result obtained by means of the primary research method; i.e., econometric analysis of impulse responses to economic crises described in Sections IIIA and IIIB of the study. While the study of impulse response functions in vector autoregressive framework is usually connected to estimating the effects of policy interventions, their analysis of the cliometric estimation indicates that their study aims at establishing (the lack of) constant event conjunction between 'negative shock at time t' and economic recovery at time $t + 10$. In other words, the study aims at analyzing whether the economies fully recover after ten years from a negative shock. The first step of the research aimed at identifying the long-term effects of recessions is to identify the timings and locations of recessions. Cerra and Saxena (2008) used several sources of narrative records to construct binary variables[13] indicating the occurrence of a crisis. Second, they estimated a univariate autoregressive[14] model with the view to estimate the average influence of the financial or political crisis on GDP growth. Third, rerunning the regression on subsamples allowed for analyzing how different groups of countries (e.g., high versus low income) respond to adverse shocks. The regression is given by the following equation (ibid., p. 441):

$$g_{it} = a_i + \sum_{j=1}^{4} \beta_j g_{i,t-j} + \sum_{s=0}^{4} \delta_S D_{i,t-S} + \varepsilon_{it}$$

Where:

g = the pace of economic development
a_i = constant for the i-th country
β_j = the autocorrelation of j-lagged g
D = digital variable denoting crisis at t
ε = error term

The average influence of a crisis on the subsequent pace of economic growth is given by δ_S. The use of linear regression and binary variables for the occurrence of various types of crises makes the model easy to interpret. Recalculating the regression for various time lags $(t-s)$ allows for estimating whether, on average, crises have a lasting effect on GDP growth. While the authors presented shock-response functions (graphs presenting average values of δ_S for $S \in < 1;10 >$, they seem to focus the discussion on the influence of crises on economic output ten years after recession took place:

> the lagged effects of currency, banking, and twin financial crises still result in 2.5%, 4%, and 5% of output loss, respectively by the end of ten years. For wars and the weakening of executive constraints, output falls initially, but at the end of ten years it is only one percentage point lower than its initial level. (p. 454)

This passage shows that Cerra and Saxena (2008) look for constant event conjunction: they use narrative records and econometric techniques to conclude that the events 'currency/banking/twin crisis' and 'economic recovery after ten years' are spatiotemporally conjoined.

The method used in the study is based on several assumptions. The authors (pp. 448–449) mention the assumption that the exclusion of feedback from previous values of GDP to the probability of a crisis and treating the occurrence of a crisis as event independent from other factors. These two assumptions seem counterfactual. At least some factors that trigger financial crises (e.g., excessive public debt) are persistent and constitute a persistent drag on GDP growth (cf. Reinhart et al. 2012). For example, the recent financial crises in Greece resulted in a sharp reduction of public spending and output, but it supposedly still lowers economic growth. In other words, the results of the study are liable to the common-cause fallacy. Both recession and slower economic recovery can be an effect of the same growth-impeding factor. Furthermore, in line with the vast majority of economic studies, the authors measure economic growth by calculating the first derivative of GDP. The authors stay silent regarding the interpretation of the error term, but, if my reading of the study as presupposing the regularity view on causality is right, then ε denotes the influence of omitted factors.

Cerra and Saxena (2008) try to control for the possibility of other causal structures, but it is still possible that both events are conjoined because of a common cause, possibly an unobservable variable. In other words, there can be a mechanism/phenomenon responsible for the co-existence of crises and later slower economic development. Therefore, the study does not deliver evidence that manipulating the cause influences the effect: counteracting the occurrence of, for instance, a financial crisis may not raise future economic development. Such an interpretation would require translating the meaning of 'causality' from the 'constant event conjunction' view into a version of the manipulationist approach that defines causality as relations invariant under interventions (cf. Chapter 6). At the same time, the studies employing the constant-conjunction view on causality deliver evidence for undertaking other kinds of actions. For example, the study definitely shows that a crisis requires undertaking contractionary steps. The following case shows that such unjustified translations are sometimes advised by authors of studies who misinterpret the evidence delivered for policymaking.

The design of the central part of Cerra and Saxena's (2008) study is aimed at analyzing the spatiotemporal connection of two events: recession and post-recessionary recovery. However, the cliometricians also conducted other analyses with a view to checking the robustness of the main conclusion and exclude the possibility of GDP growth being an endogenous variable that causes both events. To do so, they estimated several econometric models: a bivariate system of nonlinear equations (resembling vector autoregression models, cf. Section 3.2.1) and an error-correction model, among other specifications. To control for the case of reverse causality, Cerra and Saxena (2008, p. 452) collected data on growth forecasts in the years preceding crises. The data show that forecasts are too optimistic in the years before financial and banking crises and civil wars. The authors conclude based on this that crises are not predicted by economic agents and on this basis argue for their exogeneity.

Furthermore, they also estimate a probit model whereby explanatory variables are lagged values of economic growth. Dependent variables are indices for each type of crisis. They conclude that "lower (lagged) growth leads to a higher probability of crisis" (p. 454). In this case, they explicitly formulate their conclusion in line with the probabilistic condition according to which causes are probability-changing factors, cf. Chapter 3. However, this evidence plays only a supportive, confirmatory role for the main causal conclusion. In other words, the conclusion that recessions have "permanent effects" (p. 457) on economic output does not directly refer to these econometric models but results from the empirical analysis of the question whether 'negative shock at *t*' and 'economic recovery at *t* + 10' are events that are constantly conjoined. The mix of evidence supporting causal claim understood in line with regularity and probability approaches gives a hint that economists may be causal pluralists.

Finding empirical regularities

Carmen Reinhart and Kenneth Rogoff (2010) restrained from discussing the obtained results in explicitly causal terms. Their:

> main result is that whereas the link between growth and debt seems relatively weak at 'normal' debt levels, median growth rates for countries with public debt over roughly 90 percent of GDP are about one percent lower than otherwise; average (mean) growth rates are several percent lower.
>
> (p. 573)

The authors strive for employing a causality-neutral dictionary and discuss 'association,' 'link,' 'relationship,' and 'connection.' Whereas I do not agree with the statement that any of the manipulationist theories of causality are somehow privileged accounts, using causal-family words (cf. Hoover 2004) or explicitly discussing influence of one phenomenon on another seems to be a more explicit acknowledgment that one's research is of causal nature. Discussing 'regularity' or 'association' can be interpreted either as an acceptance of one of the regularity theories of causality, the logical-positivist antimetaphysical approach to science, or as an explicit refusal to interpret obtained results as evidence for causal claims. However, if the latter possibility is true, then advising intervening on the grounds of obtained results is unjustified. Therefore, Reinhart and Rogoff's policy advice may indicate that one of the former possibilities is true.

On the one hand, such formulations are in agreement with the regularity approach to causality. On the other, they are not explicit causal claims, so it is also plausible that the authors accept other, epistemically more demanding concepts of cause-and-effect relations and understand that their analysis as a purely correlational study. However, Reinhart and Rogoff (2010) seem to acknowledge the view that their results are directly applicable to economic policymaking when they ask what is the long-term macroeconomic *impact* of public debt (p. 573) or concluding their article indicating that "traditional debt management issues should be at the forefront of public policy concerns" (p. 578). Considering (1) the policy-orientation of the results and the use of 'rising'[15] in the section discussing the relation between debt, GDP growth, and inflation, and (2) the direct use of Reinhart and Rogoff's result as evidence for policymaking (e.g., Ryan 2012), I interpret their result as a causal analysis.[16]

The main conclusion of "Growth in a Time of Debt" is shedding light on the nonlinearity of relation between debt and growth (cf. Maziarz 2017; Bitar et al. 2018): "for levels of external debt in excess of 90 percent of GDP, growth rates are roughly cut in half" (p. 573). This conclusion refers to a comparatively simple quantitative approach. Namely, the authors divided their own dataset covering levels of public debt and GDP growth of 44 countries (up to 200 years) into four categories according to debt-to-GDP ratio. Each of 3,700 observations was classified into one of four categories (<30%; 30–60%; 60–90%; >90%). The next step is calculating the average pace of economic growth for each of the four categories. The use of unweighted arithmetical mean raised methodological controversies (Herndon et al. 2014), but I do not want to focus on this issue here. The obtained results suggest that countries recording debt-to-GDP ratios exceeding 90% threshold suffer from substantially lower economic development. Reinhart and Rogoff (2010) repeated calculations for different subsamples (OECD economies, emerging countries, emerged countries, shorter time series) with the view to checking the robustness of the results. Each analysis "yields remarkably similar conclusions" (Reinhart and Rogoff 2010, p. 575).

Such a research design is capable of discovering empirical regularities in datasets. Assuming that there is a mechanism or a common cause[17] that produces a nonlinear law as follows, the cliometric technique employed by Reinhart and Rogoff (2010) would yield similar results, showing that the group of observations for which public debt exceeds the 90% threshold delivers lower average GDP growth.

$$G = \begin{cases} \alpha_1 \star D + \varepsilon \middle| D < 90\% \\ \alpha_2 \star D + \varepsilon \middle| D \geq 90\% \end{cases}$$

Where:
G = the pace of economic development
D = the level of public debt
$\alpha_1;\alpha_2$ = coefficients measuring the correlation between D and G
$\alpha_1 > \alpha_2$

Accepting the logical-positivist interpretation of scientific laws (the regularity view of laws, RVL) opens the possibility to read discovered regularity in causal terms, as an instance of empirical law even without arguing for the presence of a 'necessary' connection between cause and effect. Interestingly, such threshold regularities can also be found in physics; i.e., in the discipline that was studied by the philosophers creating logical positivism. One of the examples instantiating empirical threshold laws in physics is the Wigner threshold law (cf. Chang 1970).

Despite Reinhart and Rogoff (2010) discuss the theoretical connection between public debt and economic growth, their analysis is "decidedly empirical" (p. 573), and neither informed by theory (i.e., employing aprioristic causal structure) nor used to confirm the empirical regularity found in the data. In detail, the authors mention the phenomenon of 'debt intolerance' described in their previous study (Reinhart et al. 2003), but the emphasis is put on data analysis. Considering the reductionism of the regularity view of laws (cf. Section 2.1), the causal label of the conclusion is justified, as long as one accepts the presence of an empirical regularity as sufficient evidence. Here, I should add that elsewhere (2009), Reinhart and Rogoff produced the explicitly causal claim that there is a bidirectional causal relationship between public debt and economic growth, or – to put it differently – both assertions are true that public debt causes economic growth, and economic growth causes public debt. However, in "Growth in a Time of Debt," this assertion is based not only on the search for empirical regularities, but also a theoretical conjecture about possible mechanisms producing it.

2.2.3 Other methods

Econometric and statistical methods are very fruitful at uncovering regularities in empirical data. Finding empirical regularity or constant event conjunction usually requires comparing many instances. Besides, economists use also less quantitatively advanced methods employing the methods of causal inference established by John Stuart Mill (the methods of agreement and difference) and based on Mackie's (1974) INUS condition. For example, the design of the comparative case study was used by Bard Bronnenberg et al. (2009) to study the phenomenon of 'early entry advantage.' Generally speaking, their research design was aimed at uncovering difference-making factors by analyzing datasets describing several historical characteristics of companies. They conducted several analyses of data describing consumer-goods producers with a view to identifying factors that influence market share of the companies. The use of city-level data allowed for identifying the distance to headquarters and market-entrance time as difference makers.

James Mahoney (2000) focused on studying the methods of causal inference in small-N analysis and listed the following three main approaches of (1) nominal comparison, (2) ordinal comparison, and (3) within-case analysis. The first method entails categorizing different observations into mutually exclusive and exhaustive categories. The method of nominal comparison can also be used for descriptive purposes; hence, causal conclusions should be voiced explicitly by researchers if the study is of causal nature. Such a method can be used for the verification of hypotheses about necessary and sufficient conditions. The second method resembles the method of nominal comparison but employs ordinal scales that denote rank categories that describe degrees to which a considered phenomenon is present. The ordinal comparison method resembles Mill's (2017) method of concomitant variation. I will consider the within-case analysis and process-tracing case studies in Chapter 4 because of the counterfactual nature of causal claims supported by these research methods. Julian Reiss (2009) delivered the example of qualitative comparative analysis (QCA). This method of causal inference employed by sociologists, political scientists, and economists employs the notion of causes in agreement with the INUS condition (cf. Mackie 1974). QCA is based on considering whether possibly causal factors are present across instances. The factors that are INUS conditions are labeled 'causes' of phenomena under consideration (p. 23). Case studies including within-case analysis and process-tracing design are, generally speaking, aimed at uncovering counterfactual claims about a singular event, and I will discuss them in detail in Chapter 4. However, comparative case studies, which focus on finding difference-making factors across instances, are in line with the recent developments (cf. Section 2.1.2) of the regularity theories of causality. Comparative case studies, qualitative comparative analysis, and other qualitative research designs are rarely used in contemporary mainstream economics,[18] but are more popular in management and other qualitative social sciences. However, because of the focus on economics, I do not devote much space to these research methods.

2.3 Policymaking on the basis of regularities

While I differentiate among the four different types of definitions belonging to the regularity view on causality, the distinction from the viewpoint of policymaking is twofold. On the one hand, economic research delivers causal claims based on theory (delivering causal structure) and finding empirical regularities being scientific laws. On the other hand, empirical regularities found by means of statistical techniques in observational data earn this label. In other words, the distinction is located between the results of theory-driven econometrics and the methods of finding purely empirical conjunctions. Causal claims established on the grounds of these two types of evidence have very different policy implications.

2.3.1 Cliometric results and (failed) interventions

Let me focus on the latter first. Cliometrics aims at identifying regularities in observational data. Its purpose resembles, discussed in the following chapter, the aim of atheoretical (data-driven) econometrics in excluding external (to data) knowledge from causal inference. Therefore, cliometrics share the problem of atheoretical econometrics: the causal claims produced by this type of research in principle do not warrant implementation neutrality for two reasons. First, causal claims may describe regularities produced by a common cause (common-cause fallacy). Second, the system may behave differently under intervention comparing to when it is left untouched and observed only (cf. Cartwright 2001; Dawid 2010).

Given that the cliometric studies aim at finding regularly conjoined events and do not control for other determinants of the relationship under consideration, it is possible (not to say likely) that the actual causal structure differs from what is purported by the study results. The problem is visible when one considers the result of the Reinhart and Rogoff (2010) study. The research supporting the claim that high levels of public debt are conjoined with reduced economic development does not warrant that the occurrences of these two events are not caused by external factors. For instance, both public debt and economic development may be driven by the determinants of productivity. If *A* and *B* both influence the occurrences of *X* and *Y*, but in such a way that *X* occurs before *Y*, then cliometric techniques aimed at looking for constantly conjoined events can produce false (given the true causal structure) claims. Figure 2.2 represents such a situation. Variables *X* and *Y* are driven by unknown factors *A* and *B*. The use of quantitative methods establishes regularity (correlation) of *X* and *Y* (bold arrow).

Considering that lower productivity resulting in slower economic growth or even a decline lowers government income from taxes, the observation of Reinhart and Rogoff (2010) may result from the common-cause fallacy, whereby a third variable determines both purported cause and its effect. Furthermore, it is

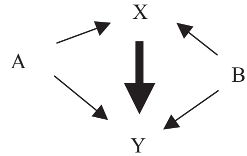

Figure 2.2 A hidden causal structure producing empirical regularity (constant conjunction) of *X* and *Y*

also possible that the direction of causality is opposite to the claim established on the basis of the cliometric analysis: slower economic development may influence government income, and therefore raise public debt. Given this, the question arises whether such purely correlational (and likely contradicting actual causal structure) knowledge is useful at all? In other words, can we act in the world when our knowledge is limited to correlational evidence that may be in disagreement with actual causal relations? Even though a skeptic would deny, I argue ahead that the answer depends on the type of intervention one considers.

My argument is based on the distinction between policy actions that do and do not break the actual (but likely unknown) causal structure. While some actions do not influence the causal structure underlying the appearances of things (e.g., the regularity uncovered by cliometrics), other interventions break causal arrows and, furthermore, require from a causal claim to be not only true but also invariant under intervention. The latter group, to be justified by evidence, require being based on causal claims describing relations invariant under intervention. However, if a causal claim is based on evidence that only allows for finding empirical regularity, but does not ascertain uncovering intervention-neutral relations, then intervening by changing the relata of such claim requires translating the causal claim from a regularity view into a manipulationist view. Without additional evidence, this translation is not justified.

Let me return to our case study of the relation between public debt and GDP growth. A typical intervention requiring a causal claim to be stable under intervention is manipulating public debt to influence economic growth. To the contrary, locating a subsidiary of a company in a less indebted country (with a view to choosing economy with a higher growth potential) would be effective (other things being equal) not only when the causal relation lacks intervention-neutrality, but also if it is spurious in the sense that the observable regularity is produced by a causal structure other than postulated by the study as long as the actual but unknown causal stays unchanged. If such a regularity is determined by other factors (since the study is based on the analysis of two types of events, it has nothing to say about the influence of other variables, and therefore, being skeptically minded, we should assume such a possibility), then our intervention on public debt may break the causal structure (arrows from *A* and *B* to *Y* at Figure 2.2), and hence, its results cannot be predicted in a reliable way. In general, the interventions that change the values of variables/features of events being relata of the causal claim used as evidence for policy require evidence asserting intervention-neutrality of causal relations. Otherwise, such interventions are based on an unjustified translation from a causal claim presupposing a regularity definition into a causal claim presupposing a manipulationist definition.

Interventions that change values of variables influence causal structure that stays unknown to us when our causal knowledge is grounded in methods aimed at uncovering causal relations understood in line with the regularity view. However, such knowledge is not useless. On the contrary, it may be fruitfully used to act in the world without influencing the causal structure underlying observed correlations. For example, a company can choose to locate its factory in a country noting lower debt-to-GDP level with a view to expecting higher profit (supposedly correlated to growth). Alternatively, a monetary union that wants its new members to have high growth potential may accept only the countries recording low levels of public debt. In such cases, the actions are unlikely to interfere with the causal structure producing regularity constituting evidence for policymaking. Other interventions require the translation of causal claims from the regularity understanding into the manipulationist notion. Taking into account that the methods of causal inference aimed at finding empirical regularities in observational data are unable to assert that these regularities hold under interventions changing the relata of the causal claim, such a translation is not justified.

To put it differently, different types of policymaking (modifying causal structure vs. using it) require different types of evidence. Unfortunately, the unjustified translations between the meanings of causal claims are considerably often in the practice of economic policymaking. For example, Reinhart and Rogoff's (2010) evidence for the regularity-view causal conclusion was indicated by numerous prominent politicians[19] as evidence for cutting government expenditure with a view to promoting economic development. Those politicians misinterpreted the causal claim as evidence that reducing debt-to-GDP ratio will improve economic development when the research method used to draw the conclusion about the association between debt and growth could only uncover the existence of empirical regularity with no warranty that interventions changing the level of debt will prove fruitful. They translated the meaning of causal claim from the regularity definition presupposed by the authors into the manipulationist notion that, unfortunately, is unwarranted by the evidence.

This kind of discussion can be compared to the widely debated thought experiment of barometer and storm. Employing the regularity definition of causality to studying the relation between barometer and weather leads to the conclusion that lower measurements of atmospheric pressure cause storms (cf. Reiss 2013; Cartwright 2007, p. 33). Given that we know the real causal structure (both barometer and weather are influenced by atmospheric pressure), the claim that barometer readings cause storms seems obviously false. However, in the case of economics and policymaking, our knowledge is often limited. While my argument could be used in support of advising getting rid of observational studies, I believe that we should use all available evidence, but keep in mind the limitations of research methods delivering the evidence we use. So, in the case of barometer and storms, while interventions breaking the causal structure (and barometer, by, e.g., raising the indicator) to save oneself from getting wet by rain are pointless, the knowledge of empirical regularities can be used to act on the basis of predictions (e.g., taking umbrellas). These examples show that even claims resulting from the common-cause fallacy can be successfully exploited for policymaking – as long as the policymaker is aware of the limitations of evidence underlying these claims.

2.3.2 Is theory-driven econometrics more reliable?

The results delivered by the project of theory-driven econometrics are, at least in principle, less susceptible to delivering causal claims being in disagreement with the actual causal structure. If the economic theory were right, causal structure imported (formally in the case of structural equation modeling and informally in the case of theory-inspired econometric models) from theoretical modeling would ascertain that empirical regularities found by means of econometrics are produced by what John Stuart Mill labeled necessary connection and later was known as scientific laws. Therefore, the response to the question put forward in the title of this subsection is, in principle, affirmative. However, the degree of certitude one ascribes to causal hypotheses based on such evidence depends on one's belief in the truth of theory delivering causal structure.

Furthermore, econometric studies, comparing to the cliometric research, deliver more reliable causal hypotheses because they control for at least some other causal factor (confounders). Hence, econometric models deliver causal hypotheses less susceptible to the common-cause fallacy or resulting from spurious correlations than the more basic cliometric techniques discussed earlier. Considering that the following Chapter 3 focuses on the discussion of establishing causal structure by means of econometric modeling, here I address the question of whether having economic theory asserting that an empirical regularity results from 'necessary connection' (i.e., is of law-like nature) is sufficient for using this type of research as evidence for interventions modifying the values of variables (i.e., imposing the manipulationist meaning on causal claims).

James Woodward (e.g., 2005, 2016) seems to be an optimist in regard to the truth of economic theory, since he interprets structural equation models as describing intervention-invariant causal relations. His epistemically courageous view on causal hypotheses delivered by structural equation modeling is depicted by the manipulationist counterfactual stating that "a change in the value of X of dX causes a change of adY" (Woodward 2016, p. 12, cf. Chapter 6). In other words, Woodward believes that the project of theory-driven econometrics is capable of delivering evidence for interventions changing the values of variables. The manipulationist interpretation can be criticized in the same way as the project of structural equation modeling: for the heavy reliance on theory (cf. Section 2.1). Ahead, I present a more skeptical stance that theory-inspired econometric models, presupposing the regularity view on causality, only justify policymaking that does not changes the values of causal relata (e.g., variables). In support, I argue that (1) economic theory can be fallible, and (2) it does not warrant causal claims to be intervention-invariant.

Let me start by discussing when Woodward's stance on structural equation models can be used as a guide for using such evidence for policymaking. If a 'theory' (i.e., [a] theoretical, axiomatic model[s] describing the causal structure underlying data under econometric analysis) is false (i.e., misses relevant causal links or delivers inexistent connections), then policy interventions suggested by a structural equation model are unwarranted. In such situations, interventions could lead to unexpected outcomes, since their actual effects would differ from those predicted. For example, a policy relying on raising average temperature to promote economic development on the grounds that sunny weather is correlated to higher profitability of stock investments (e.g., Hirshleifer and Shumway 2003) is unlikely to prove successful because it is based on a spurious correlation.

Therefore, a policymaker should consider whether the causal structure imported from theory is plausible, given the results of other studies and the strength of empirical regularity uncovered using econometrics. While all knowledge is fallible and can only be accepted provisionally, some areas of economics theorizing seem to be more trustworthy than others. Alex Broadbent (2013) offers a new method of deciding whether a new statistical result is stable or, to the contrary, will be overturned soon. This method employs consideration of whether there are other plausible hypotheses. It applies to decide if a theory is likely to be true: if there are several theoretical views on a policy setting, then the results of theory-driven econometrics can be considered unreliable. Otherwise, the result can be provisionally accepted.

Besides, even if there are a few theoretical stances inconsistently describing the same phenomena, these different theories agree concerning some causal hypotheses having policy implications. An example of a sharply divided field in economics is macroeconomics. The existence of opposing views on how the economy operates at the macro level can even be interpreted in terms of a market for ideas, where economists differentiate their theories in the process of competition (for novelty and policymakers' attention) (Scheuer and Dokurno 2017). For example, despite the fact that post-Keynesian and neoclassical macroeconomics disagree on many issues, they both predict that promoting consumption raises inflation in the long run.

Furthermore, the theory justifying causal structure must be correct in terms of depicting *the* right causal structure of phenomena without experimental closure (i.e., external influences cannot influence the policy target stronger than the determinants from the model. Otherwise, the 'manipulationist implication' of the model formulated by Woodward only holds in an experimental setting. This problem will be further discussed in Chapter 5, devoted to mechanistic causality and theoretical modeling in economics. Whether an intervention results in its effect deterministically (other things being equal), on average in an open word (i.e., not in an experiment, but in the 'field'), or on average *ceteris paribus* (other things being equal) depends on presuppositions on the nature of economic laws. Fortunately for policymaking, according to the Cowles Commission methodology discussed in Section 2.1.2, laws are interpreted as deterministic regularities, and the error term (denoting differences between actual observations and the value calculated from an econometric model) results from other factors being excluded from the regression. Therefore, if the economic theory is right and other determinants stay unchanged, then, on average, interventions should lead to the expected outcomes.

Apart from the certitude of economic theory that warrants the truth of a causal hypothesis giving evidentiary support for a policy action under consideration, a policymaker should be concerned with the problem of whether an appropriate intervention is accessible. Unfortunately for the effectiveness of economic policymaking, the right kind of 'chirurgical' intervention that changes only the value of one variable can rarely be identified beforehand, especially given the realm of social sciences. Some of the assumptions[20] of Woodward's (2005) theory of causation are unrealistic in the sense that they can rarely or never be met in an actual research/policy situations, cf. Chapter 6.

Cartwright (2001) and Dawid (2010) raised arguments in the context of directed acyclic graphs that a system under study can behave "completely different when it is kicked than when it is left alone" (p. 61). The question is if theoretical conjecture allows for closing the gap between interventional and observational regimes or if it stays open, at least in the case of social sciences. The problem of interventions changing the causal structure in a way making the policy ineffective is vividly presented by Nancy Cartwright and Jeremy Hardie (2012). The authors of *Evidence-based Policymaking* consider an intervention aiming at improving the nutrition of children in emerging countries as follows:

> educating the mother or the mother-in-law may make them feed the children better. But they may also get the idea from other members of the group where they are educated that they might get a job and hand over giving out the food to the eldest child. You have then, perhaps unintentionally, changed the social structure.
>
> (p. 31)

This example and the previous discussion of the differences in the certitude of theory-inspired econometric models show that each policy decision needs to be based on a careful study of whether a theory-driven econometric model delivers sufficient evidence to undertake action. In my view, in most cases, the translation of a causal claim produced by research methods presupposing the regularity view into the manipulationist notion that ascertains that a change in the value of some variable produces the effect as described by the causal claim is not warranted. Natural scientists seem to be more aware of the limitations of research methods. Recently, an epidemiological study showed relation ('causal' in agreement with the regularity approach) between appendectomy and lower risk of Parkinson's disease (Yeager 2018). Despite existing theoretical connection (the same proteins are found in the appendix and neurodegenerative patients' brains), the authors of the study warn that "[w]e're not saying to go out and get an appendectomy" (Neergaard 2018). The advice to wait for a piece of experimental evidence can be explained as resulting from refraining from translating the causal claim about an empirical regularity into the manipulationist notion.

2.4 You shall not translate causal claims

The regularity view on causality can either be an ontologically reductionist stance, reducing causal relations to constant conjunctions of events (or correlations), or an epistemic position according to which empirical regularities are indications of causal relations producing them. In this chapter, I have shown that even though the regularity theories are a rejected stance in the philosophy of causality debates, they still influence the research practice in economics. The influence possibly results from the chronological conjunction of the time when logical positivism accepting a version of the regularity approach to causality was a dominant stance in the philosophy of science and the development of econometrics. The use of econometric modeling inspired by theoretical conjecture is an informal version of the structural equation modeling practiced in line with the Cowles Commission methodology. In both cases, economists 'import' causal structure from theory (i.e., axiomatic models) and use econometric techniques to estimate the strength of causal relations. Cliometric techniques such as simple statistical methods or defining events using narrative records are also employed to establish constant conjunctions of events in economic history.

While these methods of causal inference differ considerably, they all presuppose a version of the regularity view on causality. To put it differently, the most epistemically courageous interpretation of what can be discovered by the research methods discussed in this chapter are either empirical regularities being 'scientific laws' or resulting from the Millean necessary connection or constant conjunction of events. Considering the fallibility of economic theory and the possibility that uncovered causal relations are spurious in the sense that they result from the common-cause fallacy or accidental correlations, these research methods do not warrant causal claims to be stable under intervention. Therefore, the interventions that change causal relata (e.g., the values of variables, features of events) are likely to be unsuccessful. First, if a causal claim constituting evidence for policymaking results from a spurious correlation, then obviously, intervention does not result in an expected change in its target. Second, the causal claims supported by theory-driven econometric models do not warrant implementation-neutrality: even if economic theory depicts the right causal structure, interventions may change it.

Consequently, translating causal claims produced by research methods presupposing a version of the regularity view into the manipulationist reading is not justified without additional evidence. Even though these arguments could be taken as opposing any policy actions based on nonexperimental evidence, my position is more modest. In real life of limited resources and epistemic limitations, one should use all available evidence. Therefore, policymakers should also use knowledge of empirical regularities. How not to misuse such evidence? The evidence on the presence of empirical regularities justifies policy actions that do not modify relata of causal claims. In other words, interventions that change the values of variables or features of events (e.g., reduce public debt) misuse causal evidence from theory-driven econometrics and cliometric studies.

Notes

1 The use of econometric techniques spans from calculating the correlation between time-series variables to estimating treatment effect while running a randomized controlled

trial. However, I believe that it is the research design and knowledge on how data were produced that matters more in the context of causal research than quantitative techniques itself. Therefore, by 'econometrics,' I mean the use of quantitative methods to studying observational data without employing experimental research design or identifying interventions.

2 I will further support this stance in the following chapters.

3 The research employing the second and third approaches will be discussed in Chapter 3 and Chapter 6, respectively. Here, I focus on discussing the differences between theoretical and atheoretical approaches to econometric modeling, and analyzing Herbert Simon's opposition to the Cowles Commission approach.

4 The question if there are empirical regularities in economics is still widely debated. Runde (1998) voiced his skepticism regarding causal explanation grounded in empirical covering laws (cf. Lawson 1997).

5 Fisher's (1922) formula states that the product of the supply of money and the velocity of money equals the product of prices and economic product ($M * V = P * Y$). Considering that the velocity of money is today measured as the ratio of GDP and money ($V = P * Y / M$) (cf. Morgan 2007), the law of the quantity theory of money is always true.

6 Because of their limited popularity in the contemporary mainstream economics, I do not exemplify them with case studies.

7 This technique is discussed in the following chapter.

8 Using own data is common among econometricians publishing in the top-quality journals. This fact can possibly be explained by the motivation of editors to draw attention (and citations). Supposedly, SEM is a framework usually employed to studying macroeconomic phenomena that may limit the number of studies based on this approach to causal inference published in the three top economic journals.

9 Ordinary Least Squares (OLS) is a simple technique of fitting a linear function into a (multidimensional) dataset so that, geometrically, the distance between observations (data points) and the linear function is minimized. If the assumptions of this estimation technique are met (cf. Wooldridge 2018, pp. 168–183), then it is an unbiased and efficient estimator.

10 Bloom et al. (2012) estimated a few regressions for different subsamples. The results are similar.

11 This statement is also true about the econometric model.

12 The hypothesis states that budgetary cuts negatively influence GDP growth in the short run. Its truth was disputed recently in the cliometric literature. It seems that the choice of the method used to identify moments of fiscal contractions determine the results (cf. Maziarz accepted).

13 A binary variable can only have two values: 0 or 1. In the study under consideration, authors constructed time-series variables for each country and each type of crisis (financial, political, and currency); 0 denotes that the crises does not occur at a given time, while 1 indicates such a crisis.

14 While the class of vector autoregression models is usually used to analyze probabilistic relations between time-series or panel data, the consideration of panel data (cf. Section 3.2), the use of this technique to analyze the influence of historical events and interpreting the results in terms of constant event conjunction indicates the presupposition of the regularity approach to causality.

15 Reinhart and Rogoff (2010, p. 575) admitted that the result "makes plain that there is no apparent pattern of *simultaneous rising* inflation and debt." (Italics come from the authors, underlying is mine.)

16 As I previously argued, "Growth in a Time of Debt," contrary to all other studies considered in the book, does not produce causal conclusions in a definitively explicit way. However, including this case serves the purpose of exemplifying how policymakers (and sometimes economists) sin by translating between different meanings of causality.

17 Here, I do not want to take sides in the ontological debate, but limit my considerations to data analysis.
18 Case studies were employed three times in articles putting forward explicitly causal conclusions published in ten years by three top economic journals (Maziarz 2018).
19 The list includes Manuel Barroso, Olli Rehn, Angela Merkel, Wolfgang Schäuble, and George Osborn (Botsch 2013; Maziarz 2017).
20 Here, I want to limit the discussion of econometrics (see Chapter 3) and the manipulationist reading of econometrics (see Chapter 6), and focus the topic of whether theory-driven econometrics do indeed produce causal claims that are invariant under interventions.

References

Adda, J., McConnell, B., & Rasul, I. (2014). Crime and the depenalization of cannabis possession: Evidence from a policing experiment. *Journal of Political Economy*, *122*(5), 1130–1202. DOI: 10.1086/676932

Alesina, A., & Ardagna, S. (2010). Large changes in fiscal policy: Taxes versus spending. *Tax Policy and the Economy*, *24*(1), 35–68. DOI: 10.1086/649828

Allen, P. G., & Morzuch, B. J. (2006). Twenty-five years of progress, problems, and conflicting evidence in econometric forecasting: What about the next 25 years? *International Journal of Forecasting*, *22*(3), 475–492. DOI: 10.1016/j.ijforecast.2006.03.003

Artuç, E., Chaudhuri, S., & McLaren, J. (2010). Trade shocks and labor adjustment: A structural empirical approach. *American Economic Review*, *100*(3), 1008–1045. DOI: 10.1257/aer.100.3.1008

Becher, E. (1905). The philosophical views of Ernst Mach. *The Philosophical Review*, *14*(5), 535–562. DOI: 10.2307/2177489

Beebee, H. (2000). The non-governing conception of laws of nature. *Philosophical and Phenomenological Research*, 571–594.

Bitar, N., Chakrabarti, A., & Zeaiter, H. (2018). Were Reinhart and Rogoff right? *International Review of Economics & Finance*, *58*, 614–620. DOI: 10.1016/j.iref.2018.07.003

Bloom, N., Sadun, R., & Van Reenen, J. (2012). The organization of firms across countries. *The Quarterly Journal of Economics*, *127*(4), 1663–1705. DOI: 10.1093/qje/qje029

Botsch, A. (2013). Hypocritical versus hippocratic economics. In Palley, Th. & Horn, G. (eds.), *Restoring Shared Prosperity. A Policy Agenda From Leading Keynesian Economists* (pp. 15–23). Düsseldorf: Boeckler.

Bottazzi, L., Da Rin, M., & Hellmann, T. (2016). The importance of trust for investment: Evidence from venture capital. *The Review of Financial Studies*, *29*(9), 2283–2318. DOI: 10.1093/rfs/hhw023

Broadbent, A. (2013). *Philosophy of Epidemiology*. London: Palgrave Macmillan.

Bronnenberg, B. J., Dhar, S. K., & Dubé, J. P. H. (2009). Brand history, geography, and the persistence of brand shares. *Journal of Political Economy*, *117*(1), 87–115. DOI: 10.1086/597301

Cameron, S., Chaudhuri, S., & McLaren, J. (2007). *Trade Shocks and Labor Adjustment: Theory* (No. w13463). National Bureau of Economic Research. DOI: 10.3386/w13463

Carnap, R. (1966). *An Introduction to the Philosophy of Science.* New York: Basic Books.

Carneiro, P., & Lee, S. (2011). Trends in quality-adjusted skill premia in the United States, 1960–2000. *American Economic Review*, *101*(6), 2309–49. DOI: 10.1257/aer.101.6.2309

Cartwright, N. (1989). *Nature's Capacities and Their Measurement*. London: Clarendon Press.

Cartwright, N. (1995). 'Ceteris paribus' laws and socio-economic machines. *The Monist*, *78*(3), 276–294. DOI: 10.5840/monist19957831

Cartwright, N. (2001). What is wrong with Bayes nets? *The Monist*, *84*(2), 242–264.

Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511618758

Cartwright, N., & Hardie, J. (2012). *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press. DOI: 10.1017/S0266267114000091

Cerra, V., & Saxena, S. C. (2008). Growth dynamics: I myth of economic recovery. *American Economic Review*, *98*(1), 439–457. DOI: 10.1257/aer.98.1.439

Chang, E. S. (1970). Study of e-H 2 scattering near the rotational threshold. *Physical Review A*, *2*(4), 1403. DOI: 10.1103/PhysRevA.2.1403

Converty, A. (2006). *Hume's Theory of Causation: Quasi-Realist Interpretation*. London: Continuum International Publishing Group.

Cooley, T. F., & LeRoy, S. F. (1985). Atheoretical macroeconometrics: A critique. *Journal of Monetary Economics*, *16*(3), 283–308. DOI: 10.1016/0304-3932(85)90038-8

Dawid, A. P. (2010). Beware of the DAG! *Proceedings of Workshop on Causality: Objectives and Assessment at NIPS 2008*, Whistler: Canada, PMLR, 6, pp. 59–86.

Dow, S. C. (2002). Interpretation: The case of David Hume. *History of Political Economy*, *34*(2), 399–420.

Dow, S. C. (2009). David Hume and modern economics. *Capitalism and Society*, *4*(1). DOI: 10.2202/1932-0213.1049

Ducasse, C. (1951). *Nature, Mind and Death*. Chicago: Open Court Publishing.

Earman, J., Roberts, J., & Smith, Sh. (2012). Ceteris paribus lost. In: Earman, J., Glymour, C. & Mitchell, S. (eds.) *Ceteris Paribus Laws* (pp. 5–26). Berlin: Springer.

Fennel, D. (2011). The error term and its interpretation in structural models in econometrics. In: Illari, Ph., Russo, F. & Williamson, J. (eds.) *Causality in the Sciences* (361–378). Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199574131.003.0017

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *222*(594–604), 309–368. DOI: 10.1098/rsta.1922.0009

Frank, Ph. (1961). *The Validation of Scientific Theories*. New York: Collier.

Frank, Ph. (2004). *Philosophy of Science: The Link between Science and Philosophy*. New York: Dover Publications, Inc.

Freedman, D. A. (1987). As others see us: A case study in path analysis. *Journal of Educational Statistics*, *12*(2), 101–128. DOI: 10.3102/10769986012002101

Gallie, R. (2013). *Thomas Reid: Ethics, Aesthetics and the Anatomy of the Self*. Berlin: Springer.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, *108*(5), 874–904. DOI: 10.1086/317671

Gelman, A., & Shalizi, C. (2012). Philosophy and the practice of Bayesian statistics in the social sciences. In: Kincaid, H. (ed.) *The Oxford Handbook of Philosophy of Social Science* (pp. 259–273). Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195392753.013.0011

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, *20*(2), 23–48. DOI: 10.1257/jep.20.2.23

Guyer, P. (1992). *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press. DOI: 10.1017/CCOL0521365872

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica: Journal of the Econometric Society*, *4*, 1–115.

Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, *51*(1), 162–172. DOI: 10.1257/jel.51.1.162

Hardt, Ł (2017). *Economics without Laws*. Cham: Palgrave Macmillan.

Hausman, D. (1983). Are there causal relations among dependent variables? *Philosophy of Science*, *50*(1), 58–81. DOI: 10.1086/289090

Hausman, D. (1998). *Causal Asymmetries*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511663710

Hempel, C. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.

Henschen, T. (2018). What is macroeconomic causality? *Journal of Economic Methodology*, *25*(1), 1–20. DOI: 10.1080/1350178X.2017.1407435

Herndon, T., Ash, M., & Pollin, R. (2014). Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. *Cambridge Journal of Economics*, *38*(2), 257–279. DOI: 10.1093/cje/bet075

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, *58*(3), 1009–1032. DOI: 10.1111/1540-6261.00556

Hitchcock, Ch. (2010). Probabilistic causation. In: Zelta, E. (ed.) *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/win2016/entries/causation-probabilistic/. Access: 21st March 2010.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. Thousand Oaks, California: SAGE Publications.

Hoover, K. D. (2001). *Causality in Macroeconomics*. Cambridge: Cambridge University Press.

Hoover, K. D. (2004). Lost causes. *Journal of the History of Economic Thought*, *26*(2), 149–164. DOI: 10.1080/1042771042000219000

Hoover, K. D. (2008). Causality in economics and econometrics. In: Durlauf, S. & Blume, L. (eds.) *The New Palgrave Dictionary of Economics*. New York: Palgrave Macmillan.

Howitt, P. (2010). Endogenous growth theory. In: Durlauf, S. & Blume, L. (eds.) *Economic Growth* (pp. 68–74). London: Palgrave Macmillan.

Hume, D. (1956). *A Treatise of Human Nature by David Hume*. Gloucestershire: Clarendon Press.

Hume, D. (1992). *An Enquiry Concerning Human Understanding*. London: Hackett Publishing.

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2006). Toward a comprehensive developmental model for major depression in men. *American Journal of Psychiatry*, *163*(1), 115–124. DOI: 10.1176/appi.ajp.159.7.1133

Keuzenkamp, H. (2000). *Probability, Econometrics and Truth: The Methodology of Econometrics*. Cambridge: Cambridge University Press.

Lawson, T. (1997). *Economics and Reality*. London: Taylor and Francis.

Leigh, D., Devries, P., Freedman, C., Guajardo, J., Laxton, D., & Pescatori, A. (2010). Will it hurt? Macroeconomic effects of fiscal consolidation. *World Economic Outlook*, *93*, 124.

LeRoy, S. F. (2018). Implementation-neutral causation in structural models. *Contemporary Economics*, *12*(3), 253–267.

Loptson, P. (1998). Hume, multiperspectival pluralism, and authorial voice. *Hume Studies*, *24*(2), 313–334.

Mach, E. (1960). *The Science of Mechanics*, translated by Thomas J. McCormack. La Salle, IL: Open Court, pp. 157–182.

Mackie, J. L. (1974). *Cement of the Universe*. Oxford: Clarendon Press. DOI: 10.1111/j.1468-0149.1975

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly*, *2*(4), 245–264.

Mahoney, J. (2000). Strategies of causal inference in small-N analysis. *Sociological Methods & Research*, *28*(4), 387–424. DOI: 10.1177/0049124100028004001

Maziarz, M. (2015). A review of the Granger-causality fallacy. *The Journal of Philosophical Economics: Reflections on Economic and Social Issues*, *8*(2), 86–105.

Maziarz, M., & Mróz, R. (2019). Response to Henschen: Causal pluralism in macroeconomics. *Journal of Economic Methodology*, First View. DOI: 10.1080/1350178X.2019.1675897

Maziarz, M. (2017). The Reinhart-Rogoff controversy as an instance of the 'emerging contrary result' phenomenon. *Journal of Economic Methodology*, *24*(3), 213–225. DOI: 10.1080/1350178X.2017.1302598

Maziarz, M. (2018). Causal inferences in the contemporary economics. *Mendeley Data*. Retrieved from: http://dx.doi.org/10.17632/v7dhjnd8xg.2. Access: 16th October 2018.

Maziarz, M. (accepted). It's all in the eye of beholder. *Argumenta Oeconomica*.

Mill, J. S. (2017). *The Collected Works of John Stuart Mill*. Wien: e-artnow.

Moneta, A. (2005). Causality in macroeconometrics: Some considerations about reductionism and realism. *Journal of Economic Methodology*, *12*(3), 433–453. DOI: 10.1080/13501780500223742

Morgan, M. S. (2007). An analytical history of measuring practices: The case of velocities of money. In: Boumans, M. (ed.) *Measurement in Economics: A Handbook* (pp. 105–132). Bingley: Emerald Insight.

Neergaard, L. (2018). Appendix removal is linked to lower risk of Parkinson's. *The Washington Post*. Retrieved from: www.washingtonpost.com/national/health-science/appendix-removal-is-linked-to-lower-parkinsons-risk/2018/10/31/3cd4bdaa-dd37-11e8-8bac-bfe01fcdc3a6_story.html?noredirect=on&utm_term=.91f910038813. Access: 1st November 2018.

Pap, A. (1952). Philosophical analysis, translation schemas, and the regularity theory of causation. *The Journal of Philosophy*, *49*(21), 657–666. DOI: 10.2307/2020991

Pearl, J. (2012). The causal foundations of structural equation modeling. In: Hoyle, R. (ed.) *Handbook of Structural Equation Modeling* (pp. 68–91). New York: Guilford Press.

Psillos, S. (2009). Regularity theories. In: Beebee, H., Hitchcock, Ch. & Menzies, P. (eds.) *The Oxford Handbook of Causation* (pp. 133–156). Oxford: Oxford University Press.

Reichenbach, H. (1971). *The Theory of Probability*. Berkley and Los Angeles: University of California Press.

Reichenbach, H. (2012a). *The Direction of Time*. Berkley: University of California Press.

Reichenbach, H. (2012b). *Hans Reichenbach: Selected Writings 1909–1953*, Vol. 2. Dordrecht: Reidel Publishing Company.

Reid, Th. (1815). *The Works of Thomas Reid*. Edinburgh: Samuel Etheridge Publishing.

Reinhart, C. M., Reinhart, V. R., & Rogoff, K. S. (2012). Public debt overhangs: Advanced-economy episodes since 1800. *Journal of Economic Perspectives*, *26*(3), 69–86. DOI: 10.1257/jep.26.3.69

Reinhart, C. M., & Rogoff, K. S. (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton: Princeton University Press. DOI: 10.3386/w13882

Reinhart, C. M., & Rogoff, K. S. (2010). Growth in a time of debt. *American Economic Review*, *100*(2), 573–578. DOI: 10.1257/aer.100.2.573

Reinhart, C. M., Rogoff, K. S., & Savastano, M. A. (2003). *Addicted to Dollars* (No. w10015). National Bureau of Economic Research. DOI: 10.3386/w10015

Reiss, J. (2009). Causation in the social sciences: Evidence, inference, and purpose. *Philosophy of the Social Sciences*, *39*(1), 20–40. DOI: 10.1177/0048393108328150

Reiss, J. (2013). Contextualising causation part I. *Philosophy Compass*, *8*(11), 1066–1075. https://doi.org/10.1111/phc3.12074

Rodrik, D. (2015). *Economics Rules: The Rights and Wrongs of the Dismal Science*. WW Norton & Company. DOI: 10.17323/1726-3247-2015-4-39-59

Romer, C. D., & Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review*, *100*(3), 763–801. DOI: 10.1257/aer.100.3.763

Runde, J. (1998). Assessing causal economic explanations. *Oxford Economic Papers*, *50*(2), 151–172. DOI: 10.1093/oxfordjournals.oep.a028639

Ryan, P. (2012). *The path to prosperity: A blueprint for American renewal.* New York: Macmillan.
Salmon, W. (1989). *Four Decades of Scientific Explanation.* Pittsburgh: University of Pittsburgh Press.
Scheuer, B., & Dokurno, Z. (2017). Introduction: Philosophical foundations of modern economics. In: Dokurno, Z. (ed.) *The Paradigms of Contemporary Macroeconomics from the Perspective of Sustainable Development.* Warsaw: PWN.
Schlick, M. (1936). Meaning and verification. *The Philosophical Review*, *45*(4), 339–369.
Shipley, B. (2016). *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R.* Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511605949
Simon, H. (1957 [1977]). Causal ordering and identity. In: Simon, H. (ed.) *Models of Man: Social and Rational: Mathematical Essays on Rational Human Behavior in Society Setting.* (pp. 55–80). Dordrecht: Springer.
Simon, H. A. (1951). A formal theory of the employment relationship. *Econometrica: Journal of the Econometric Society*, 293–305. DOI: 10.2307/1906815
Sims, Ch. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48. DOI: 10.2307/1912017
Stöltzner, M. (2009). The logical empiricists. In: Beebee, H., Hitchcock, Ch. & Menzies, P. (eds.) *The Oxford Handbook of Causation* (pp. 108–127). Oxford: Oxford University Press.
Strawson, G. (2014). *The Secret Connexion: Causation, Realism, and David Hume: Revised Edition.* Oxford: Oxford University Press.
Stroud, B. (2000). 'Gliding or staining' the world with 'sentiments' and phantasms'. In: Rupert, R. & Richman, K. (eds.) *The New Hume Debate* (pp. 16–30). London: Routledge.
Suppes, P. (1970). *A Probabilistic Theory of Causality.* Amsterdam: North-Holland Publishing Company.
Tabellini, G. (2010). Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, *8*(4), 677–716. DOI: 10.1111/j.1542-4774.2010.tb00537.x
Weber, M. (1992 [1905]). *Die protestantische Ethik und der 'Geist' des Kapitalismus.* Düsseldorf: Verlag Wirtschaft und Finanzen.
Whiteley, P. F. (2000). Economic growth and social capital. *Political Studies*, *48*(3), 443–466. DOI: 10.1111/1467-9248.00269
Woodward, J. (2005 [2003]). *Making Things Happen: A Theory of Causal Explanation.* Oxford: Oxford University Press.
Woodward, J. (2016). Causation and manipulability. In: Zelta, E. (ed.) *The Stanford Encyclopedia of Philosophy.* Retrieved from: https://plato.stanford.edu/archives/win2016/entries/causation-mani/
Wooldridge, J. M. (2018). *Introductory Econometrics: A Modern Approach*, 5th International Edition. Boston: Cengage Learning.
Wright, J. (1973). *The Sceptical Realism of David Hume.* Manchester: Manchester University Press.
Wright, J. (2000). Hume's causal realism: Recovering a traditional interpretation. In: Read, J. & Richman, K. (eds.) *The New Hume Debate* (pp. 88–99). London: Routledge.
Yeager, A. (2018). The neurodegenerative disease shares protein clumps in common with appendixes, perhaps explaining why removing the organ is protective. *The Scientist.* Retrieved from: www.the-scientist.com/news-opinion/appendectomy-may-lower-risk-of-parkinsons-disease-65026?utm_content=79231847. Access: 1st November 2018.
Ziliak, S., & McCloskey, D. N. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives.* University of Michigan Press. DOI: 10.3998/mpub.186351

3 Causality as changes in conditional probability

Employing the regularity account of causality is problematic if scientists face reality without constant regularities. The development of quantum physics and the rapid growth of the social sciences in the second half of the twentieth century created a demand for a philosophical theory of causality that can deal with stochastic relations. In response to this need, Hans Reichenbach, Irving Good, Norbert Wiener, Nancy Cartwright, Brian Skyrms, Ellery Eells, and others formulated their versions of the probabilistic definition of causality. All the views included in the probabilistic approach share the belief that the appearance of a cause raises the probability that its effect will happen or that causes modify the probability distribution of their effects. In general, the probabilistic theories of causality are prevalent in defining the relation in the following way: C causes E if and only if (1) C and E are spatiotemporally connected, and (2) the occurrence of cause C raises the probability that effect E will occur: $P(E \mid C) > P(E)$.[1] (cf. Section 3.1) The probabilistic approach to causality inspired the development of several methods of causal inference. The most straightforward applications are Granger-causality tests accompanied by other techniques belonging to the set of atheoretical econometrics and the method of Bayesian nets (directed acyclic graphs, or DAGs) that is of limited popularity in economics.[2] Section 3.2 exemplifies the methods of causal inference allowing for concluding in terms of changes in conditional probability with cases of recent economic research: the use of the vector-autoregressive model as a method of time-series analysis and estimating a cross-sectional model without importing causal structure from theory. In Section 3.3, I argue that atheoretical econometrics do not warrant causal claims to be invariant under intervention, and I analyze a historical case of monetary policy based on the St. Louis equation as an example of intervening on the basis of spurious causal claims.

3.1 Probabilistic theories of causality

Section 3.1 summarizes the probabilistic approach to causality. First, I discuss the most influential probabilistic theories of causality in chronological order. Second, I analyze some issues that are relevant to the study of presuppositions implicitly accepted by economists. Third, I review the debates present in the philosophy of economics related to the probabilistic view on causality and analyze whether Granger causality deserves the causal label. Finally, I address the question of whether the commonsense claim that correlation does not imply causation is true.

3.1.1 The menu of probabilistic definitions

Reichenbach's (2012b [1956]) book published posthumously is chronologically the first analysis of causality in terms of probability. Contrary to his first attempt devoted exclusively to epistemic considerations (cf. Chapter 1), in his later work, Reichenbach focused on developing an ontological concept of time. With a view to developing the asymmetry of the fourth dimension, he studied causality but, contrary to other philosophers of causality, he did not presuppose the time precedence of the cause but wanted the asymmetry to result from the study of causality. He formally defined causality in terms of time precedence and a change in probability: "[i]f probability implication is valid in only one direction, then the antecedent is the temporally later event" (Reichenbach 2012b, p. 94). The account can be formalized as follows: $(C \rightsquigarrow E) \wedge\neg (E \rightsquigarrow C)$. Here, $\rightsquigarrow$ means 'implies with probability'. A similar conclusion indicating the inferential asymmetry was formulated by Baumgartner (2008). Furthermore, Reichenbach (2012a) disagreed with a causal interpretation of functional equations on the grounds that they lack the asymmetry, which, according to his belief, is a crucial feature of causality: "[f]unctional relationship . . . is a symmetrical relation; if y is a function of x, then x is a function of y" (p. 28). Reichenbach's significant contribution is the development of the principle of common cause. Events A and B are caused by C if they occur together more often than they would if they were independent: $P(AB) > P(A)\mathrm{P}(B)$ (Reichenbach 2012a, p. 157 et seq.). Additionally, his definition of screening off inspired the Bayesian nets approach (cf. Pearl and Verma 1995) to causal inference.

Norbert Wiener (1956) was interested in problems arising from predicting time series. His work does not seem to use the causal label systematically, but Granger and the philosophers working on the methods of causal inference based on studying partial correlations later employed his views (cf. Section 3.1.3). Interestingly, a similar change in language can be observed in the works of Judea Pearl, who, at the beginning of his work on inferential methods, (e.g., 1982) discussed inference from data staying agnostic about the nature of conclusions while, later, explicitly discussed causality (e.g., Pearl and Verma 1995).

In detail, Wiener (1956) defined causal relation between two continuous functions in the following way: "[f]or two simultaneously measured signals, if we can predict the first signal better by using the past information from the second one than by using the information without it, then we call the second signal causal to the first one" (p. 127). In other words, S_1 causes S_2 if predictions based on a model containing all the relevant information and past observations of S_1 are more accurate than those produced by a model excluding S_1. This type-level definition, which is the first concept binding causally functions (variables) instead of events, inspired Clive Granger to develop his econometric tests (Granger 1980). The distinction between Wiener's and Granger's views on *prima facie* causality between variables is that the former definition focuses on continuous functions, while the latter on discrete variables. Wiener's conception was previously extended to the discrete time-series analysis by William Stahlman (1948) in his doctoral dissertation, but Granger was probably unaware of his results.

Patrick Suppes (1970) developed his theory of probabilistic causality and opposed identifying such relations with constant regularities because "the everyday concept of causality is not sharply deterministic in character" (p. 7). By discussing a few case studies (e.g., reckless driving raising the chance of an accident), he justified the view that the 'causality as probability raising' approach is especially crucial for the social sciences. On the grounds of a few thought experiments, Suppes (1970, p. 9) stated that "[t]he omission of probability considerations is perhaps the single greatest weakness in Hume's famous analysis of causality." He coined the following definition of causality:

> one event is the cause of another if the appearance of the first event is followed with a high probability by the appearance of the second, and there is no third event that we can use to factor out the probability relationship between the first and second events.
>
> (p. 10)

In other words, C causes E if $P(E\,|\,C) > P(E)$ and C and E are not produced by the same cause (the requirement of the lack of common cause).

In his theory of probabilistic causality, Suppes (1970) distinguished between *prima facie* (naïve) causality that only considers the likelihood of occurrence of two events (i.e., other causal factors are excluded from the analysis) and the fully fledged concept that takes into consideration all relevant factors. *Primafacie* causality proved later to be crucial for Granger-causality tests and some other econometric techniques discussed in Section 3.2. Suppes defined *primafacie* causality as follows: "[t]he event $B_{t'}$ is a prima facie cause of the event A_t, if and only if (i) $t' < t$, (ii) $P(B_{t'}) > 0$, $P(A_t \mid B_{t'}) > P(A_t)$" (p. 12). In other words, event $B_{t'}$ causes event A_t only if it precedes its consequent in time, the probability that B_t occurs is above zero and – crucially – a cause raises a probability of its effect. Interestingly, Suppes (1970, pp. 48–52) believed that his probabilistic approach to causality could be employed to an analysis of singular events. Such an approach would require understanding probability as a subjective belief in likelihoods (cf. Keynes 2013) instead of employing the frequency account usually used in empirical scientists.

Suppes' (1970) theory of causality is ambitious regarding the ontology of causality. On the one hand, the unspecific considerations of mechanisms indicate that Suppes belongs to the objectivist camp. On the other hand, his acknowledgment of mental causation indicates an antirealist flavor. As Williamson (2009, p. 192) put it, Suppes "is a pluralist about causality – causality varies according to the interpretation of probability, the conceptual framework, and the notion of mechanism under consideration." Despite the plurality of Suppes' views, the theory put forward in his notable book (1970) is a standard formulation of the probabilistic approach to causality. Suppes' definition of causality proved crucial for several contemporary methods of causal inferences. Suppes' views are presupposed in Granger-causality tests, VAR modeling, and the graph-theoretic approach.

Contrary to Suppes, Irving Good (1959), in his earlier account, did not employ the indices of time to make his account plausible with the backward causation that seems to exist in the realm of quantum mechanics. His later (Good 1961a, 1961b) attempts at measuring the strength of causal tendency was later employed to the Bayesian nets framework. The measure is given by $log \frac{P\left(\bar{E}|\bar{F}H\right)}{P(\bar{E}|FH}$; i.e., the logarithm of a change in the probability that an effect occurs. Good was a causal realist and treated the probabilistic measure of the strength of causal tendency as an epistemic concept instead of a reduction of causality to changes in conditional probability.

Nancy Cartwright (1979) developed Suppes' (1970) concept of *primafacie* causality with the aim of distinguishing between spurious causality (when other factors influence the probabilities) and the real causal effects. She (p. 423) produced the following definition: "C causes E if and only if C increases the probability of E in every situation which is otherwise causally homogeneous with respect to E" what can be formalized as $P(E|CK) > P(E|K)$ where K denotes all the other causes of E (cf. Williamson 2009, p. 193; Cartwright 2007, p. 45).

Clive Granger (1969), widely known in econometrics and other data-intensive sciences for the causality tests bearing his name, is one of the few econometricians that put forward an explicit definition. Even though economists usually refer to his concept as 'Granger causality' to differentiate the results of these econometric tests from an 'actual' causality (cf. Section 3.1.3), his definitions belong to the philosophical tradition of the probabilistic approach. In detail, Granger (2012) acknowledged that his method accepts Hume's viewpoint, but also Wiener's (1956) inspiration is clearly visible. Hoover (2008, p. 12) indicated that the Granger-causality tests exemplify the modern, probabilistic approach to cause-and-effect relationships. According to Granger (1980), causal relations are characterized by (1) time precedence, (2) invariance in time,[3] and (3) including additional information about the effect. Putting it differently, a cause precedes its effect in time, the relation does not change in time, and adding a cause to a model improves the quality of predictions of its effect. The general definition of causality put forth by Granger (1969) states that Y_n causes X_{n+1} if:

$$P(X_{t+1} \in A \,|\, \Omega_n) \neq P(X_{t+1} \in A \,|\, \Omega_n - Y_n)$$

Where:
A = a real number
$X_t; Y_t$ = two time-series variables
Ω_t = all the knowledge available in the universe on t

A direct application of such a definition of causality is impossible for two reasons, at least. First, obtaining Ω_t; i.e., all the past knowledge available in the universe, is undoubtedly epistemically impossible. Second, even if such knowledge were in principle accessible, such a model would be impossible to estimate conclusively because the number of variables would exceed the number of instances so that the number of degrees of freedom would be lower than zero. In order to solve this issue, Granger used Suppes' (1970) idea of *prima facie* definition and Wiener's (1956) concept of causality as predictability. If X influences the probability distribution of Y, then employing the past information on X to predicting future values of Y will improve accuracy. In other words, if X and Y are *prima facie* Granger-causally connected, then a model containing all past observations of X and Y is, regarding the accuracy of predictions, superior to a model containing exclusively the past of X. Formally, *Yprima facie* Granger-causes X if:

$$P\left(X_{t+1} \in A | X,\ Y\right) \neq P(X_{t+1} \in A \mid X)$$

Where:
A = a real number
$X; Y$ = time series
X_{t+1} = the value of X on $t+1$

On the one hand, the definitions such as Granger's general definition that contain 'all relevant information,' 'all causal factors,' and other synonymical phrases cannot be directly applied in research practice without prior knowledge about other causally relevant determinants. While such definitions can serve as ontological views, they are not suitable as an epistemic stance serving the purpose of research guidance. On the other, the application of *prima facie* definitions coined by Wiener, Suppes, and Granger can often lead to obtaining spurious results. With a view to solving this issue, Skyrms (1980) argued that it is justified to conclude that C causes E if C raises the probability of E at least in one context. This can be formalized as follows: $P(E \mid X_i \ \& \ C) > P(E \mid X_i \& \sim C)$. Accepting this weaker definition allows for putting forward causal conclusion, even though C brings about E only in some contexts. As I will argue ahead, economists seem to accept a variation of this definition in their research. The 'context unanimity' of a definition of causality seems to be the right solution instead of choosing between possibly misleading *prima facie* definition and metaphysical assumption of considering all relevant causal factors. However, John Dupré (1984) disagreed on the grounds that such a definition is susceptible to criticism similar to the counterarguments against the versions of *prima facie* causality: the inclusion of all relevant factors to a causal model is never guaranteed.

I have reviewed the most relevant theories of causality belonging to the probabilistic approach with a view to present the menu of the probabilistic definitions. In the following section, I consider some differences between these theories. Ahead, I focus on the philosophy of economics debates related to the probabilistic approach.

3.1.2 Criticism and further differences

All the versions of the probabilistic theories of causality presented previously have in common accepting that causes raise the probability of their effects or that effects appear more often in the presence of their causes. They differ regarding ontological views, defining what is 'probability' discussed in the definiens of causality, indicating the relata of causal relations, the context in which a cause affects the probability of its effect, and some more issues. I am far from trying to take sides in these debates, and limit the purpose of this section to presenting the pluralism of the probabilistic approach with a view to offering a background to the reconstruction of presuppositions held by economists using the methods of causal inference presupposing a version of the probabilistic approach to causality.

All of the probabilistic theories discussed earlier define causal relations in the following form: *X* causes *Y* if and only if a probability measure of *X* is influenced by *Y*. Such definitions of causality can be interpreted in two ways. On the one hand, the probabilistic view can be taken as an ontological stance; i.e., as a reduction of causal relations to changes in conditional probability (cf. Williamson 2009, p. 187). In other words, a probabilistic theory of causality can assert that there is nothing in causal relations beyond changes in conditional probabilities. An excellent example of taking the probabilistic theory as an ontological view is Reichenbach's position. According to his stance, causality is nothing more than (different from) changes in conditional probability. Reichenbach (2012b) put forward the probabilistic theory of causality with the hope for reducing causal language to probabilities and hence allowing for causal modeling in the sciences in an objective manner.

On the other hand, the probabilistic notion of causality can be interpreted as an epistemic theory. In such a case, one can stay agnostic about an ontological account of causality or accept a different stance as depicting what causality is and take the probabilistic definition as guidance in research. The only explicit definition of causality voiced on the grounds of econometrics seems to be such a theory. While Granger (1969), in earlier writings, did not take sides on this issue, he later (1980) explicitly ascribed to the viewpoint that Granger-causality tests should be used as a source of evidence supporting causal hypotheses instead of being an ontological view on what actual causality is. Regarding the ontology of causality, Granger seems to be a realist when he admitted that his tests for causality should not be used without any theoretical support for the presence of causal hypotheses (ibid.).

The distinction between the epistemic and ontological interpretations of the probabilistic theories of causality determines the view on probabilities. If the probabilistic definition is accepted as a metaphysical stance, then the probabilities should stand for an objective feature of phenomena. In such a case, a right candidate for the stance on probabilities is either a frequentist interpretation assuming that, in the limit, the frequencies of re-run experiment indicate the actual probability of the objective, underlying process, or, in line with Skyrms' (1980) view, taking probabilities as a rational belief. On the contrary, if the probabilistic approach is taken as an epistemic concept, then probabilities can be taken as a within-model representation of chances (cf. Suppes 1970) or as a subjective belief (if one considers the case of actual [singular] causation).

Generally speaking, there are two views on what are the relata of causal relations. Sometimes causal relations are interpreted as binding singular events/facts/'particulars' (token causation). Othertimes, causality is a relation between types of (groups of) singular instances (type causation) such as variables, types of events, etc. The proponents of different theories of causality had different views on what the relata are. As Ehring (2009, p. 387) admitted, in the probabilistic framework, "[t]he most popular candidates are events and facts, but there are alternatives including tropes, exemplifications of universals, and objects." Events are usually defined as spatiotemporally located single occurrences of particular situations that happen only once, and therefore measuring its probability in line with the frequentist paradigm is impossible. If such a definition were accepted, then probabilities had to be defined in terms of subjective belief (e.g., in line with Keynes' (2013 [1921]) view presented in his *A Treatise on Probability*). Therefore, if a version of this presupposition is rejected, then the relata are defined as classes of considerably similar events. This allows for using the most popular (frequency) interpretation of probabilities in causal inference. Other relata of token-level causal claims are the values of variables (Spirtes et al. 2012) or variables (Henschen 2018). These two views are most widespread in the sciences relying on quantitative modeling. Following, I argue that econometricians take variables or types of events as the relata (cf. Section 3.2). Arguing for a more general conclusion that variables instantiate properties of objects and the properties are causal relata seems convincing (Hausman 1998, p. 87). Accepting both types of relata or views on causality (type and token levels) is also possible. For example, Eells (1991) distinguished between type-level causality relating types of events and token-level holding between events and argued that both levels are separate: considering the probabilistic nature of causal relations, the presence of causality at the type level does not assert that an instance of a generalization; i.e., token-level causal relation is always affected by the type-level relation. Apart from discussing what causal relata are, the connection between causal relationships is sometimes considered (Ehring 2009). Do causal relations hold only between two events, or is causality a multi-relational concept? In the case of the probabilistic approach, two views can be acknowledged. The use of *prima facie* definition states that causality holds between a cause and its effect only, and hence accepts causal relations to be binary. However, the more developed definitions take causality to be a relation among several causal factors and their effect.

Even though the probabilistic theories evolved from the regularity approach in response to accepting statistical laws in the sciences half a century ago, the question if a cause is required to raise the probability of its effect always, in some contexts (at least one) or the majority of situations waits for a conclusive answer. All views can be found among the probabilistic theories of causality discussed earlier. While Suppes (1970) has not addressed this issue explicitly, his theory suggests that he believed the causes of events to influence the conditional probability across all contexts (in all situations). This requirement was liberalized by Skyrms (1980), who argued that influencing the chance that an effect appears in at least one causal context is sufficient. Accepting a *prima facie* definition equates to presupposing the requirement that a cause should raise the probability of its effect *on average* across contexts specific to research (e.g., in a given sample). Assuming that researchers' data come from different contexts, then estimating the change of conditional probability required by the probabilistic definition delivers the positive result (i.e., indicating causal relation) only if the cause under consideration raises the probability of the effect more (not to be mistaken with 'more often') across all contexts covered by the database used in a study. The latter two responses seem to be presupposed in econometric research that is based either on testing for *prima facie* causality or a more advanced definition referred to by Granger as 'general.' Depending on the type of data used in a study, econometricians conclude that two variables are causally related on the grounds of studying one or several contexts. The solution to the dispute seems to differ depending on the purpose of one's definition. If one strives for a definition of causality that can be used in actual scientific research, then Skyrms' solution seems to be a good ground for building econometric tests of causality. Furthermore, the pragmatic considerations taking into account what are the effects of committing type I and type II errors (i.e., taking an unrelated factor as a cause and rejecting an actual cause) can play a role in deciding which of the requirements should be presupposed for particular research.

Is probabilistic causality evidence for actual causality?

Generally speaking, the main criticism of the probabilistic theories of causality focuses on highlighting the distinction between spurious changes in probabilities that are not causal and actual causal relations. Two strains of literature can be distinguished. On the one hand, some scholars highlight that the definitions of causality supported by different versions of the probabilistic view are not sufficient for distinguishing between actual causal relations and the correlations resulting from a common cause or spurious correlations. In fact, all the methods discussed in the following section are based on the assumption that accidental, non-causal relations will disappear in the limit (when the number of observations increases infinitely) and analyzing infinite samples is impossible. However, this assumption is unverifiable and sometimes fallacious (Nowak 1960). An example of this way of criticism is Hesslow's (1976) argument against Suppes' (1970) version of the probabilistic theories of causality. On the basis of considering Suppes' *prima facie* definition, Hesslow concluded that some events that fulfill this version of the probabilistic causality are not genuine causes. A similar problem was observed by Edward Simpson (1951), who coined a counterexample (the Simpson paradox) indicating that actual causes, in the presence of confounding factors, can lower the probability instead of raising it. Wesley Salmon (1980) criticized the previously discussed Good's theory of probabilistic causality for mistaking causality with a statistical association and argued that Good's measure of strength designates the strength of a statistical association instead of a causal relation. This way of criticizing the probabilistic approach usually proceeds by delivering thought experiments that are counterexamples in the sense that the actual causal structure depicted by a story differs from the one obtained by causal inference grounded in a version of the probabilistic definition.

On the other hand, philosophers supporting other families of theories criticize the probabilistic approach for the impossibility of inference about how a studied system behaves under intervention. The methods of causal inference aimed at uncovering probabilistic dependencies in the data such as Bayesian nets (DAGs) and econometric models (cf. Section 3.2) estimate probabilities by a quantitative study of observational data. Because interventions usually break some causal relations (and sometimes are assumed to break them all and set values of some variables, cf. Chapter 6), predicting the effects of interventions on the grounds of observational studies may be impossible. Nancy Cartwright (2001) and Philip Dawid (2010) raised such arguments against causal inference with Bayesian nets (DAGs). However, similar reasoning applies to marcoeconometric models, and, by analogy, other quantitative studies of purely observational data. In response to such criticism, Judea Pearl (2009, pp. 85–89) developed a framework for coping with calculating probabilities from actual interventions.

The probabilistic approach to causality does indeed have some drawbacks that are appropriately underlined by philosophers opposing such theories of causality. Nevertheless, at least some of the counterarguments miss the point of why the probabilistic definitions have been formulated and gained the audience of quantitatively oriented scientists. Let me consider the case of Granger-causality tests that are heavily criticized on similar grounds (e.g., Hoover 2001). The purpose of developing Granger-causality tests was not to offer a single and error-free method of causal inference, but to deliver a tool (one among many) for studying economic phenomena. Granger (1969, 1980) himself seems to be aware of the limitations of his definition of and tests for causality when he advised using the tests only for testing causal hypotheses deduced from theories. The use of the Granger-causality tests as the only causal evidence makes the results lack informative conclusions due to possible errors (Maziarz 2015). In a similar vein, many counterarguments are only valid in the context of the thought experiment, while the actual research practice proves that the probabilistic view on causality is useful for some purposes, even though it may lead to erroneous inferences.

3.1.3 Probabilistic causality in the philosophy of economics

The philosophers of economics and methodologically oriented economists seem to be skeptical about the probabilistic view on causality. In the philosophy of economics, considerable attention of the disputes related to this approach to causality is put on addressing the question if the quantitative studies of observational data are sufficient evidence for causal claims or, to put it differently, whether probabilistic causality is the actual causality.[4] Those philosophers and economists who accept the probabilistic view on causality usually refrain from deeper ontological considerations and address particular questions arising from obstacles faced by researchers in their day-to-day practice. An example of such a problem is a philosophical analysis of the problem of spurious correlations or the issue of choosing model specification in the atheoretical econometrics.

However, there are a few exceptions. Alessio Moneta (2005) interpreted the probabilistic approach to causality as an antirealist and reductionist position and exemplified his view with the case of the Granger causality tests. Interpreting all the probabilistic theories as being reductive is not justified in the light of the earlier discussion of the multitude of theories within this approach. At least some of the probabilistic theories of causality are epistemic in nature and refrain from delivering definite opinions on the ontology of causality. Therefore, it is possible to take probabilistic evidence as an indication of the presence of causality understood in line with other (possibly realist) definitions.

To support his view, Moneta (2005) repeated the accusations of some econometricians (e.g., Leamer 1985) who disagree with the causal interpretation of the results of Granger-causality tests on the grounds that such tests are an indication of predictability rather than causality. Econometricians sometimes also voice their skepticism regarding this method of causal inference. For example, Maddala and Lahiri (2009, p. 390) admitted that Granger causality is not "causality as it is usually understood." Indeed, the presence of Granger causality does not imply the exogeneity of variables in the structural-equation framework (cf. Chapter 2). Neither the existence of exogeneity indicates the presence of Granger-causality (cf. Hoover 2001, pp. 151–155). Another problem with this concept of causality is connected to how testing for Granger causality proceeds. For example, there are many tests for causality that do not deliver the same results for the same data. The subjectivity is also connected to choosing the number of lags (cf. Section 3.2) that can overturn previous results. However, similar problems are connected to other methods of econometric modeling, so the issue at stake seems not to be connected to the causal label, but rather to methodological problems that also apply to other econometric techniques.

On the contrary, others rightly indicate that the definition put forward by Granger (1969) instantiates the probabilistic view and should be interpreted as an epistemic, not ontological, concept (see Granger 1988). For example, Cartwright (2007, p. 29) admitted that this approach to testing for causality agrees with Suppes' (1970) definition. Both stances employ the requirements of temporal connection and probability raising. The difference worth noting is that while Suppes was interested in causal relations between events, the relata of Granger causality are variables. In a similar vein, Hoover (2001, p. 150), despite supporting a version of the manipulationist definition himself, defended this approach as follows:

> Granger-causality is a species of the probabilistic approach and subject to all the general objections that can be leveled against it. We cannot, however, dismiss it as an unsuccessful account of causality on those grounds alone. On the one hand, the probabilistic relations that constitute Granger-causality may yield important information for causal judgments even if they do not

> adequately define them – that is, they may have an informational part to play in causal inference, even if they are not constitutive of causation.

In this passage, Hoover agreed that Granger-causality tests could be used instrumentally, even though one rejects the probabilistic definitions of causality as misdescribing what causality really is. Therefore, those philosophers and econometricians who criticize the method of testing for Granger causality (e.g., on the grounds that Granger causality diverges from what is believed to constitute causal relations, cf. Henschen 2018, p. 2) in time series seem to disagree with the probabilistic view in general rather than with the particular definition offered by Granger.

Nevertheless, the methodological criticism of the Granger causality made econometricians cautious concerning formulating conclusions in causal terms. They (e.g., Chang et al. (2014) usually conclude that variable *A* "Granger-causes" variable *B* instead of merely labeling the relation causal. Moneta (2005, p. 442) explained this linguistic tradition by conjecturing that econometricians support a manipulationist view on causality and "[t]he concept of Granger-causality is much weaker than controllability." Nevertheless, the popularity of this method of causal inference has continuously grown in the last decades. Granger-causality tests have been employed beyond economics. Moneta (2005, p. 436) admitted that the Granger-causality testing "is maybe the most influential procedure of causal inference in econometrics." The tests are also used by neuroscientists, biologists, ecologists, and epidemiologists (Maziarz 2015).

In some cases, economists also take part in philosophical debates. An example of a philosophically minded economist is John Hicks (1979), who aimed at analyzing the similarities and dissimilarities between the concept of causality in physics and economics. One of the differences is that, in economics, the demand for time precedence is not, according to Hicks, always fulfilled: cause and effect can happen at the same time. Considering that some of the present-day physicists accept even backward causation (i.e., the situation whereby an effect precedes its cause in time [Cramer 1986]), such a viewpoint does not seem to be counterintuitive. Nevertheless, the backward causation, being counterintuitive, is still not a widely accepted stance (Hardt 2017, p. 101). Another example of the philosophical orientation of econometricians is Robert Engle et al.'s (1983) definition of causality. The econometricians argued that causes in an econometric model are the exogenous variables that determine their effects (the variables on the left-hand sides of equations), and the conditional distribution is invariant under interventions (the requirement is labeled superexogeneity). This definition belongs to the family of the manipulationist approach. While the requirement of super-exogeneity can serve as a normative definition, I argue in Section 3.3 that econometric studies (without additional knowledge on the phenomena under study) never warrant it to be fulfilled.

Despite the few attempts of the philosophers of economics to develop views on causality and take part in the debate on the 'big' philosophical questions, they usually focus on actual methodological problems faced by econometricians.

An example of such a debate is the exchange of views between Kevin Hoover (2003) and Julian Reiss (2007) who disagreed on whether cointegrated variables are correlated. Despite being in disagreement, both arguments are framed as responding to Elliot Sober's (2001) counterargument against the principle of common cause. Roughly speaking, PCC, formulated by Reichenbach (2012a [1956]), states that correlated variables have common causes. PCC underlies the quantitative techniques used to draw causal conclusions from observational data such as directed-acyclic graphs (DAGs). According to Sober's counterargument, the principle is contradicted by the case of sea level in Venice and bread prices in the United Kingdom: both variables have been rising for the last centuries and therefore are correlated. However, they obviously are not causally connected.

Hoover (2003) defended the principle of the common cause by arguing that the two variables discussed by Sober are, in fact, not correlated: they are cointegrated. Cointegration denotes the situation whereby two variables are trending in time (are nonstationary) in the same direction. In such a situation, econometric textbooks advise differentiating variables to make them stationary. Stationarity, contrary to nonstationarity, describes the situation when the average value of subsamples of time series is constant. The correlation resulting from the nonstationarity of time series is spurious and disappears when the variables are detrended. Reiss (2007) disagreed with both Hoover (2003) and Sober (2001), and claimed (1) that PCC should be interpreted as delivering evidence for the presence of causality that is not a deterministic law, but rather an empirical regularity and therefore can be fallacious in some cases, and (2) that the samples of cointegrated variables are indeed correlated. Regarding the former claim, Reiss (2007) argued that PCC should not be considered as an ontological concept but only as a tool for delivering empirical evidence for causal claims: it "cannot serve as a metaphysical principle in a definition of causation. . . . It is rather used as an epistemic principle for causal inference." As a solution, Reiss advises reformulating the principle of the common cause by adding the requirement of suitable preparation (i.e., preprocessing the data in accordance with the methodology of econometric modeling) of the variables. According to Reiss (2007, p. 11, emphasis in original), the PCC should be defined as follows: "[i]f two *suitably prepared* random variables *X*, *Y* are probabilistically dependent, then either *X* causes *Y*, *Y* causes *X* or *X* and *Y* are the joint effects of a common cause *Z*."

The dispute seems to be driven by having in mind different purposes of calculating correlations between variables. It is true that calculating the correlation of two cointegrated variables leads to the conclusions that they are correlated. Furthermore, they do have a common cause, *time*, as long as this dimension of empirical data is accepted as a causal factor. However, it is also true that the estimated correlation may lead us astray if it is taken as an indication of a causal connection for policy purposes. Therefore, if a study strives for delivering causal evidence, preprocessing data by calculating a derivative of time series can help in dealing with spurious causality. However, if a study is used for the prediction of one variable with the help of another variable, then detrending may not improve the fit to data (more on this in Section 3.3).

3.2 Testing for probabilistic dependencies

Generally speaking, the probabilistic approach to causality won the hearts of the quantitatively oriented economists. According to the distinction between econometric approaches presupposing the regularity and the probabilistic view on causality put forth in Chapter 2, economists who (implicitly) accept a version of the probabilistic stance employ data-driven and atheoretical techniques. All techniques share in common use of the knowledge about time precedence to establish the direction of causality between correlated variables. The Granger-causality tests, directly employing one of the probabilistic definitions, are an obvious example. In contemporary economics, the tests are rarely employed as an exclusive method of causal inference probably because the results of Granger-causality tests used as exclusive evidence can lead researchers astray (cf. Maziarz 2015). In Section 3.2.1, I analyze the analysis of Jaeger and Paserman (2008) that studied the Israeli-Palestinian conflict by means of testing for Granger causality in a vector-autoregressive framework. The case is a prime example of drawing causal conclusions from time-series data.

However, economists also arrive at causal conclusions understood in agreement with a version of the probabilistic definitions by studying cross-sectional data. In the case of datasets that do not consist of time-indexed variables, the atheoretical grounds for distinguishing between causes preceding effects in time are less obvious (cf. Maziarz 2018). In general, economists use knowledge of how data were generated. For example, Schechter (2007) took as causes the values of the variables that were collected before the variable interpreted as the effect). A different method is to consider the chronological order of phenomena for which the variables stand. For instance, Lefgren et al. (2012) considered variables describing father's wealth and child's welfare, and established the direction of causal relationships on the grounds of extra-economics knowledge that child's welfare follows in time father's income. In Section 3.2.2, I analyze Stock et al.'s (2006) study of the causes of dropouts from Ph.D. programs in economics that instantiate such an approach and establish the direction of causality on the factual knowledge on the time precedence of phenomena represented by the variables included into the analysis.

Apart from the econometric techniques that belong to the repertoire of contemporary economics, there are other methods such as data-mining algorithms and methods developed within the Bayesian nets framework that are rarely used in contemporary economics. They also allow for putting forward causal claims presupposing a version of the probabilistic approach. Considering that these methods are not, despite their popularity in other sciences, used in mainstream economics (Maziarz 2018), I refrain from discussing them. Numerous philosophers attempt to interpret these techniques in line with a version of the manipulationist approach to philosophy (cf. Section 6.1). I oppose this view for two reasons. First, the techniques of Bayesian nets have been inspired by the probabilistic theories of causality developed by Hans Reichenbach and Irving Good. Second, if Bayesian nets and other data-mining techniques are employed to studying observational data, the same criticism that I formulate in Section 3.3 against taking econometric evidence as indication of relations invariant under interventions applies. Therefore, the manipulationist conclusions are not justified.

3.2.1 Causal inference from time-series data

The use of time-series data seems to be the distinct feature of econometrics compared to other data-intensive sciences. The quantitative methods employed by researchers from other disciplines usually utilize cross-sectional or panel data that describe the features of some populations. On the contrary, econometricians (especially those interested in the macroeconomic phenomena) are often interested in interdependencies between two variables representing the development of a phenomenon under study in time. For example, macroeconometricians study the relation between the prices of bread in Britain and the level of the sea in Venice or, what is more popular, money and inflation. Frequently, the study of two variables cannot be informed by economic theory because it is either underdeveloped or split (Maziarz 2019) and, therefore, observational data are the only available evidence.

An example of such a study is the analysis of David Jaeger and Daniele Paserman (2008). The two econometricians analyzed the relation between the death-resulting violence in the Palestinian-Israeli conflict. While the authors start their paper from discussing Thomas Schelling's (1960) theoretical analysis of conflict, the paragraph-long summary at the beginning of the introduction serves the purpose of catching readers' interest rather than informing the structure of their econometric estimation.[5] On the grounds of studying exclusively two variables denoting the number of deaths inflicted by Israeli and Palestinian forces, the authors concluded that "one Palestinian fatality raises the cumulative number of Israeli fatalities by 0.25 [. . . and] one Israeli fatality raises the number of Palestinian fatalities by 2.19" (p. 1603). Jaeger and Paserman used an explicitly causal language when admitting that the Israeli responses to Palestinian attacks "*caused*" ten times more fatalities than the Israeli tit-for-tat violence (ibid.). The causal language is, in fact, present from the beginning of the study, where the authors admit that "[o]ur primary interest is the *effect* of 'own' fatalities on fatalities of the *opposite* group" (p. 1594) (italics are original, the underlying mine). In other words, Jeager and Paserman are interested in the influence of the aggression suffered by one side of the conflict on the violent response of that side. They concluded that "the Israelis react . . . to Palestinian violence against them, while Palestinian actions appear not to be related to Israeli violence" (2008, p. 1603). The causal claim obviously refers to the only research method used by the authors, which is testing for the presence of Granger causality in a vector-autoregressive framework.[6]

Generally speaking, the VAR models explain the present values of variables by their past. The VAR approach dates back to the work of Christopher Sims

(1980) who advocated for practicing econometric modeling in an atheoretical way and opposed the Cowles Commission approach advising informing the structure of econometric models from axiomatic theory; i.e., choosing exogenous variables in an aprioristic way (cf. Canova 1999). In the historical context, the set of VAR models consists of reduced-form and structural models. "A *reduced form* VAR expresses each variable as a linear function of its own past values, the past values of all other variables being considered and a serially uncorrelated error term" (Stock and Watson 2001, p. 103) (italics are original, the underlying mine). A two-variable, first-order (i.e., containing only one lag of the variables) VAR model (Verbeek 2012, p. 351) can be described by the following equations:

$$X_t = \delta_1 + \theta_{11} Y_{t-1} + \theta_{12} X_{t-1} + \varepsilon_{1t}$$

$$Y_t = \delta_1 + \theta_{11} Y_{t-1} + \theta_{12} X_{t-1} + \varepsilon_{2t}$$

Where:
$Y_t ; X_t$ = the values of time-series variables at t
δ_1 = a constant
θ_{11} = the parameter denoting one-period autocorrelation of Y
θ_{12} = the parameter denoting the influence of X_{t-1} on Y_t
ε_{1t}; ε_{1t} = error terms

The structural-form (SVAR) models are derivable from the reduced-form equation by adding additional (theory-driven) restrictions, which is referred to as the process of identification. The identification is aimed at introducing the causal structure among contemporaneous variables. As Demiralp and Hoover (2003, p. 747) indicated, economists believe that:

> no empirical or statistical basis for the choice of the contemporaneous causal orderings [exists . . . , so that economists must appeal to a priori knowledge. . . . Practitioners typically regard the members of the family [of SVARs implied by a reduced-form VAR] as observationally equivalent.

Although VAR models, in general, are an example of an atheoretical method of inference, the structural VAR models are liable to the same accusations of aprioristicness as the Cowles Commission methodology of structural modeling (Moneta 2005).

Historically, the two types of VAR models have been introduced with the view to analyzing the macroeconomic phenomena and solving to the problem of underdeveloped or inconsistent macroeconomic theory. Today, the reduced-form[7] VAR models belong to the standard repertoire of methods used for time-series analysis and are also used beyond the domain of macroeconomics. These models are usually employed for the purpose of predicting the values of interrelated variables. In fact, the VARs have become popular for their predictive power, which is superior in comparison to the structural equations. Despite the fact that the 'father' of the VAR framework (Sims 1980, p. 12) read the reduced-form VARs in causal terms, this interpretation is criticized by several economists and methodologists (e.g., Zellner 1988; Stock and Watson 2001; Demiralp and Hoover 2003). The opponents argue that only the structural models justify formulating conclusions in a causal language. Nevertheless, the case under consideration shows that, under some conditions, economists formulate causal claims on the grounds of reduced-form VARs.

The authors of the case study paper assumed that the empirical reaction functions have the form of the following VAR regression:

$$\begin{pmatrix} Pal_t \\ Isr_t \end{pmatrix} = A_0 + A_1 \begin{pmatrix} Pal_{t-1} \\ Isr_{t-1} \end{pmatrix} + \ldots + A_{p/q} \begin{pmatrix} Pal_{t-p} \\ Isr_{t-q} \end{pmatrix} + \boldsymbol{BX_t} + \varepsilon_t$$

Where:

Pal_t = the number of Israeli fatalities (i.e., the Palestinian response at *t*)

Isr_t = the number of Palestinian fatalities (i.e., the Israeli response at *t*)

$A_{p/q}$ = the vector of coefficients denoting the influence of $\begin{pmatrix} Pal_{t-p} \\ Isr_{t-q} \end{pmatrix}$ on the current values

$\boldsymbol{BX_t}$ = the influence of other determinants (day of the week; the chronological phase of conflict; and the length of the barrier on the West Bank border)

ε_t = the error term

The method serves the purpose of testing "whether fatalities on one side of the conflict cause fatalities on the other side" (p. 1594). Unfortunately, Jaeger and Paserman (2008) remained silent on their approach to choosing the model specification (the other variables included in the model X and the number of lags p/q). The estimation of reduced-form VARs (vector autoregressive models) cannot be informed by an aprioristic knowledge on the causal structure or the length of lags (t–1 in this example). Therefore, the process of model estimation resembles, to some degree, the data-mining approach and is based on choosing one of several regressions estimated in the process. The estimation usually proceeds on the equation-by-equation basis by means of OLS (ordinary least squares).[8]

Economists favor the following two ways of choosing the number of lags (p) in the model. First, one can preprocess the data using a *prima facie* Granger-causality test with the aim at discovering dependencies between lagged explaining variables and an explained variable: only those lags (p) are included that significantly Granger-cause the explained variable (cf. Verbeek 2012, p. 351 et seq.). Another common approach is to estimate a model including all variables covered by a dataset of the researchers and then exclude some variables until the quality of predictions is insufficient. A version of this approach is to use the information criteria (such as Akaike's Information Criterion AIC or Schwartz's Bayesian Information Criterion SBC) (Lütkepohl 1990). Contrary to these popular approaches, Jaeger and Paserman (2008) used a third approach and estimated the average reaction functions and their statistical significance for a few aprioristically chosen[9] values of lags (p/q). Notice that failing at specifying an appropriate length of the lagged values leads to obtaining spurious results (cf. Maziarz 2015 for a review).

The second step of Jaeger and Paserman's (2008) study is to test for the presence of Granger causality in the estimated VAR models. To do so, the authors employed the so-called Granger direct test,[10] which is the first test for Granger causality offered by Granger himself (Granger 1969). The original test is based on the *prima facie* version of Granger's definition and happens to be one of the most straightforward tests for Granger causality. Let me assume for simplicity that the test is only used to test for the presence of unidirectional causality $X \rightarrow Y$. In this case, Granger direct test is conducted as follows. First, the regression specified following is estimated. Second, the F-statistic is employed to decide whether the parameters a_{2_m} differ significantly from 0. Rejecting the null hypothesis of the test[11] leads to the conclusion that X Granger-causes Y. The statistical significance of a_{1_n} indicates that Y is autocorrelated.

$$Y_t = a_0 + a_{1n} \sum_{n=1}^{N} \mathrm{Y}_{\mathrm{t-n}} + a_{2m} \sum_{m=1}^{M} \mathrm{X}_{\mathrm{t-m}} + \varepsilon_{\mathrm{t}}$$

The genuine version of the Granger direct test is based on the *prima facie* definition. On the contrary, Jaeger and Paserman (2008) included an additional vector X of variables. There are many econometric tests of Granger causality: e.g., Granger Direct Test, General Granger-causality Test, Sims Test, Modified Sims Test (Granger 1969; Sims 1972). The majority of econometric tests locates themselves in between the two definitions put forward by Granger (*prima facie* and general) and indicate that X causes Y when X has additional information useful for predicting Y that is not contained in a previously identified set of causally relevant variables. In other words, for practical purposes, Ω from the general definition denotes a set of causally relevant variables instead of a set of 'all information available in the universe'. These variables are usually picked up on the basis of a priori knowledge (theoretical models) or previous econometric studies.

Additionally, testing for causality obviously cannot employ the whole history of a variable. The number of lagged values (n; m) is chosen considering theoretical knowledge about the process (e.g., for monthly data, including 12th lagged value may improve the accuracy of predictions) or practical reasons such as sample size. The definition of causality put forth by Granger can be used in more advanced tools of time-series analysis such as VARMA models (Lütkepohl 2005). Jaeger and Paserman's testing procedure is certainly not based on the *prima facie* definition. Including day of the week, war period, and the length of the separation barrier is the operationalization of 'all knowledge' present in

Granger's (1969) general definition. In other words, these three additional variables are taken as relevant correlates of the number of fatalities in the conflict. The obtained result made Jaeger and Paserman (2008) conclude that "Palestinian violence Granger-causes Israeli violence" (p. 1594).

It is interesting to observe the movement from the 'Granger causality' conclusion to the (general) causal claim just cited. The economists (p. 1954) are aware of the difficulties connected to interpreting the results of Granger-causality tests as the evidence for 'true' causal claims. Unfortunately, Jaeger and Paserman (2008) have never specified what they mean by the 'true' causality. While a version of the manipulationist view would be an intuitive candidate, the evidence delivered in support of their conclusion does not warrant that you could manipulate the number of Israeli fatalities to influence the level of their aggression on Palestinians. Similarly, Jaeger and Paserman (2008) are not entitled to produce their claim in terms of mechanism due to the lack of a theoretical study of the phenomenon. Therefore, it seems that they accept, at least as an epistemic stance, a version of the fully fledged probabilistic definition (i.e., not the *prima facie* definition). My interpretation is further supported by the discussion of past correlations and the low likelihood of the presence of the common-cause fallacy.

To proceed from Granger causality to 'true' causality, Jaeger and Paserman (2008, p. 1594) attempted to refute the possibilities that a third variable causes both variables under consideration, or that the direction of causality has an opposite sign. I should highlight here that even the more advanced tests that are based on Granger causality instead of the *prima facie* concept are liable to the misspecification of what information is 'relevant'; i.e., including inappropriate variables in the set of relevant information excluding X (Ω). The problem of common-cause fallacy has interestingly been exemplified by Atukeren (2008), who showed that chocolate Easter bunnies *prima facie* Granger-cause Easter. Including additional variables (i.e., appropriate implementation of Granger's general definition) leads to the conclusion that such a relation is spurious.

Furthermore, considering that the testing procedure involves estimating regression and verifying whether some parameters significantly differ from zero, all the issues usually discussed in the context of econometric modeling apply. Uncontrolled cointegration can result in obtaining a significant false result, but differentiating time series can also produce spurious results (Lee et al. 2002). Nonlinear causal relations are usually undetected due to testing for a linear relationship. Additionally, modifying sampling frequency can change or even reverse the direction of Granger causality, producing an unintuitive result (McCrorie and Chambers 2006).

Nevertheless, despite these issues, the results are useful for some purposes. The evidence delivered by testing for the presence of Granger causality in a VAR framework does not allow for concluding that the manipulation of a cause will influence its effect, but can be used for predicting or policymaking that does not involve influencing the causal structure of phenomenon under consideration. I will develop this view in the following Section 3.3. What are the X and Y, or, specifically, what are the relata of causal claims supported by the study under consideration?

The causal conclusion put forward by Jaeger and Paserman (2008) holds between two variables denoting the number of fatalities in the Israeli-Palestinian conflict that were compiled by the authors on the grounds of statistics published by B'Tselem, an NGO studying human rights violations at the territory of the conflict. Sometimes, variables are interpreted as being quantitative descriptions of properties of phenomena or objects (Hausman 1998). However, in the case of the studied econometric analysis, the two variables seem to stand for events defined as 'at day *t*, *X* Palestinians/Israelis were killed in the conflict.' Accepting such an interpretation allows for stating that the primary causal conclusion of the paper holds between types of events.[12] However, running Granger-causality tests in the VAR model of a macroeconomic phenomenon (e.g., the variables representing money and inflation) seems to presuppose the variables as features of phenomena, since the event interpretation would not make much sense in such a context.

One of the topics debated in the philosophy of causality literature devoted to the probabilistic approach (cf. Section 3.1) is the question of when causal relations should produce observable results. The answers range from the least requiring view that causes must raise the probability of their effects at least in one context, through standard responses that it must be the case in several or most contexts, to the answer that causes must always raise the likelihood that their effects will happen. The last answer is definitely of ontological nature: it is in principle impossible to verify that some type of event or phenomenon influences the probability of its effect across all contexts. Jaeger and Paserman's (2008) claim is spatiotemporally limited to one region and historical context (post-WWII). The economists do strive for producing a generalization describing the nature of all conflicts, but limit their endeavor to that particular case. Therefore, it seems that they acknowledge the least demanding requirement for causes, according to which they raise the probability of their effects in at least one context.

Another topic considered in the philosophical debate is the interpretation of probabilities. While Jaeger and Paserman (2008) remained silent on the issue, their study seems to presuppose the frequency reading. In general, stating that *X* causes *Y* on the grounds of a Granger-causality test in the VAR framework refers to the situation where variable *X* changes conditional distributions of *Y*. Considering the atheoretical nature of this approach to practicing econometrics, the conclusion that the distributions are inferred empirically is obvious. The frequency interpretation of probabilities follows from the empirically inferred distribution of variables. Of course, statistical tests are aprioristic in the sense that they are (practically always) based on the Gaussian curve, even though economic variables are often described by other distributions.

Jaeger and Paserman's study establishes the direction of causality on the grounds of time precedence. In the case of using time-series variables, this step is easy and straightforward. However, econometricians sometimes use cross-sectional data and, what I show in the next subsection, draw causal conclusions from econometric studies also presupposing a version of the probabilistic view on causality.

3.2.2 Atheoretical, cross-sectional models

In the previous chapter, I argued that the cross-sectional regressions, in which structure is either derived from or inspired (in an informal way) by theory, presuppose the regularity view on causality that defines such relations as empirical regularities instantiating laws of nature. In such cases, economic theory delivers evidence in favor of the presence of a 'theoretical connection' that supports the claim that the regularity uncovered by means of quantitative study is not accidental. However, in some areas of econometric research, economic theory is either split or underdeveloped (Maziarz 2019), making informing the structure of the econometric model from economic theory impossible. Wendy Stock et al. (2006) faced such a situation when they studied the factors that make doctoral students in economics quit. Their article starts by indicating that "[r]emarkably little is known about the timing and extent of attrition of doctoral students from economics Ph.D. programs" (p. 458). Therefore, neither theoretical knowledge from economics, sociology, and psychology, nor empirical studies that remain silent on this particular topic,[13] allow for informing the structure of the econometric model.

Cross-sectional and panel datasets can be compared to a frame from a movie and a movie. While the former describes the features of some population at a particular moment in time, panel data show how these features change over time. What follows, cross-sectional data, are not explicitly indexed in time. Considering that partial correlations are symmetrical relations, there is no direct way to establish the direction of causality on the grounds of time precedence when studying cross-sectional data. To do so, econometricians need to consider what phenomena are quantitatively represented by variables and use extra-economic knowledge. For example, Almond et al. (2005) studied econometrically a dataset covering the relationship between birth weight and the level of income at adulthood. Even though narrowly understood data cannot be used exclusively to hypothesize both strength and direction of a causal relation, conjecturing that one's birth happens before adulthood (as long as the premise excluding backward causation is acknowledged) allows for hypothesizing that the causal relation $\textit{birth weight} \xrightarrow{raises} \textit{income}$ produces a positive correlation.

However, Stock et al. (2006) decided to import the research design from epidemiology, where it is known as prospective (sometimes cohort) study. In medicine, data scientists gather data by choosing a population and recording exposure to a risk factor under consideration (e.g., smoking) and outcomes (e.g., lung cancer). This research design is superior to using observational data gathered post-factum (retrospective study) for at least two reasons. First, it allows for studying the development of phenomena in time, so that even cross-sectional data could constitute causal evidence grounded in time precedence without any background knowledge. Second, the prospective study methodology solves the problem of instances that fall out of the sample. In the case of an epidemiological study, the individuals who died because of lung cancer are unlikely to respond to a questionnaire invitation. In the case of Ph.D. studies attrition, analyzing the population of 'survivals' (i.e., the students that have not resigned until some point) may strongly bias the estimates.

Considering these issues, Stock et al. (2006) decided to follow a cohort of 586 doctoral students who entered Ph.D. programs in economics in fall 2002 (p. 458).[14] The econometricians gathered several variables characterizing the Ph.D. programs such as quality estimated by the National Research Council, faculty-student ratio, whether written exams are required after the first year, whether students have access to office space at the campus, etc. They also questionnaired the participating students about their demographic characteristics and prior educational and academic attainment (pp. 458–459). This preliminary dataset was used as independent variables (determinants or causes) in the model described ahead. The dependent variable (i.e., the effect) denoting whether students finished the first and second years of their studies was delivered by the institutions. As an additional, exploratory study, Stock et al. (2006) repeated the questionnaire, asking those Ph.D. students who resigned about their career goals and reasons for the decision. The use of a sample of the population of economics Ph.D. students indicates that the causal claims hold not only in the context of the analyzed subpopulation, but are aimed at describing the whole population of economics Ph.D. students in the United States.

On the grounds of the prospective study, the econometricians identified several *reasons* for attrition. These include the quality of the Ph.D. program and previous academic attainment, but also access to shared office space. The determinants of dropout rates are established on the grounds of a standard probit model specified as follows[15] (p. 462):

$$\Pr(dropout_i = 1) = \phi\left(\beta_0 + \beta_1 \star P + \beta_2 \star S\right)$$

Where:
ϕ = normal cumulative density function
β_0 = a constant
$\beta_1; \beta_2$ = coefficients
P = Ph.D. program characteristics
S = Ph.D. student characteristics

The model delivers an estimation of the probability that student *i* will drop out from their Ph.D. program. The explanatory variables are divided into program characteristics (*P*) and student characteristics (*S*), while the endogenous variable is a binary (being 1 for students who resigned from their doctoral program and 0 elsewhere).

The probabilistic nature of the specified relation results from the features of probit models. The label refers to a class of models whereby the dependent variable has only two values: 0 and 1. There are several methods of estimation,[16] such as maximum likelihood, Berkson's minimum chi-square method, and some newer, computationally more demanding, techniques (e.g., Gibbs sampling) (cf. Borooah 2002, p. 45 et seq.). The model calculates the probability of attrition for each student, given the characteristics *P* and *S*. Generally speaking, it is a two-step procedure. First, a linear function of all determinants is calculated as a weighted average ($\beta_0 + \beta_1 \star P + \beta_2 \star S$). Second, the obtained result is transformed so that it fits the cumulative normal distribution. The overall result is interpreted as the probability that, for a given set of values of independent variables (*P*, *S*), the independent variable equals 1. If a causal interpretation is imposed on the model, then the probit estimation can be read as follows: particular values of independent variables determine the probability that the independent variable equals 1 is *p*. Considering that probit models are estimated in a way that, roughly speaking, aim at minimizing the number of wrongly classified instances, the frequency view on the nature of probability seems to be implicitly accepted by the econometricians what bears resemblance with the previous case study.

The direction of causal relation results from the time precedence of explanatory variables. Both the characteristics of doctoral programs *P* and doctoral students *S* stand for events and phenomena that have happened and have been collected before Ph.D. students resign (or are forced to leave, in some cases). Therefore, on the premise that causes precede their effects in time, the econometricians can segregate their dataset into the explanatory and explained variables even though the partial correlation[17] is a symmetrical relation. In other words, unless the assumption of time precedence is introduced, another causal structure (e.g., suggesting that dropout rates cause the size of a Ph.D. program) is equally plausible.

Considering that (1) the probit estimation employed by Stock et al. (2006) delivers the estimates of how particular sets of parameter values (variables having specified values) determine the probability of dropout, and (2) causes are prior in time to their effects, the model under consideration discovers probabilistic relations in the dataset – and therefore a version of the probabilistic definition of causality seems to be presupposed by the researchers. Otherwise, discussing the 'reasons' for attrition (p. 463) or association (p. 459) would not be justified. On the grounds of the probit model, Stock et al. (2006) formulate general (type-level) causal claims about the influence of some characteristics of program or students on attrition. In the previous subsection, I argued that the variables describing the number of fatalities in the Israeli–Palestinian conflict stand for event types. While the endogenous variable is a quantitative (digital) description of the event (dropout of an *i*-th student), interpreting the exogenous variables is open to more than one plausible reading.

On the one hand, each of the exogenous variables can be taken as a quantitative description of features of a phenomenon (e.g., the size of a Ph.D. program).

On the other hand, a holistic look at the model and the set of explanatory variables suggests that all these variables characterize events understood as enrollment of *i*-th *S*-like student at *P*-like program. If the latter interpretation is accepted, the probit model estimated by Stock et al. (2006) can be read as evidence for the causal claim that has types of events as its relata. In the case of the former interpretation, a set of factors (instead of a single type of event), such as characteristics of phenomena characterized by variables, are causal relata.

Given this, the definition of causality presupposed by Stock et al. (2006) would hold between types of events (enrollment of *S*-like student at *P*-like program) determine (with approximation/stochastically) the probability that another event type (attrition from a Ph.D. program) occurs. In detail, the probit model, as interpreted by Stock et al. (2006) seems to presuppose the following definition of causality: C causes E if the probability of E conditional on C differs from the unconditional distribution $P(E \mid C) \neq P(E)$. If my reconstruction is correct, then econometricians diverge from the standard (e.g., Suppes 1970) view that causes must raise the probability of their effects and prefer the broader notion of causes modifying (i.e., raising or lowering) probabilities of their effects.

In summary, the probabilistic view on causality is presupposed in econometric research, whereby econometricians attempt at delivering evidence supporting causal claims by a quantitative study of observational data in an atheoretical way. Data-driven econometric analysis of time-series or cross-sectional data allow for establishing causal claims that are based on a version of the probabilistic definition that resembles Granger's general definition and Cartwright's (1979) definition whereby 'all knowledge' and 'all other factors influencing an effect' are operationalized by researchers that choose a limited number of other relevant factors. The relata of such causal claims are types of events and features of phenomena described quantitatively by variables.

Ahead, I argue that the causal claims established on the grounds of atheoretical econometrics (within an n-dimensional model) may not be true in an $n + m$–dimensional policy setting. Therefore, such evidence does not warrant interventions modifying the relata of causal claims to be successful. Nevertheless, the knowledge on probabilistic dependencies can be successfully employed to policy actions that do not require causal claims to be invariant under intervention.

3.3 The common-cause fallacy and policy actions

The econometrics studies previously discussed implicitly presuppose a version of probabilistic causality according to which causes raise the probability of their effects, given some background factors. Such a definition is located between the simple concept of *prima facie* causality (that takes into consideration only the probabilities of two events or conditional distributions of two variables) and the fully fledged definition (requiring knowledge of *all* relevant factors). The rejection of the philosophically more sound view requiring causes to raise the probability of events given all relevant factors results from the epistemic limitations.

First, the problem is that all relevant causal factors of the event under consideration are unobservable – and some of them remain unknown. Second, even if all relevant background conditions could be identified in practice, econometric estimation requires having a higher number of observations than the number of variables (the number of the degrees of freedom needs to exceed 1), and hence only a selected set of variables can be included to a regression.[18] Therefore, while the inclusion of all causally relevant factors allows for defending the probabilistic definition from accusations of accepting spurious correlations as causal relations, such rich definitions of probabilistic causality cannot be directly applied to causal inference.

To put it differently, Cartwright's (1979) definition,[19] Wiener's (1956) concept of signal predictability, and Granger's general definition can serve as ontological views on what causality really is, but these fully fledged conceptions are impossible to test econometrically, and therefore a less demanding definition is employed to the research practice.[20] The need to limit the number of background conditions (variables) in econometric models makes causal claims susceptible to the common-cause fallacy. This label denotes a situation when the observed (spurious) relationship between C and E results from both C and E being caused by another variable (cf. Figure 3.1). The frequency of causal claims based on data-driven econometrics resulting from the common-cause fallacy stays unknown, but the problem is likely to occur considerably often, given that econometricians face the phenomenon known as multicollinearity (many economic variables change in time, and therefore are correlated to each other).

Helmut Lütkepohl (1982) argued that the existence of common-cause fallacy leads to the following inferential asymmetry. Let me assume that an econometrician tests for Granger causality in a bivariate process (i.e., employs the concept of *prima facie* Granger causality to study the relation between two variables). In such a situation, rejecting the null hypothesis (i.e., finding Granger causality) can only be interpreted as evidence for causality in this bivariate process but does not indicate the presence of Granger causality in a higher-dimensional process. This is the case since the causal evidence from a bivariate process says nothing about the influence of other factors. On the contrary, accepting the null (finding no evidence for Granger causality) can be extrapolated onto a higher-dimensional process. In other words, *prima facie* Granger non-causality implies Granger non-causality, but *prima facie* Granger causality does not imply Granger causality.

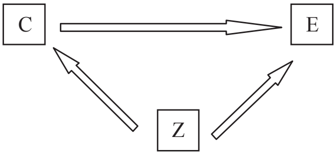

Figure 3.1 The common-cause fallacy

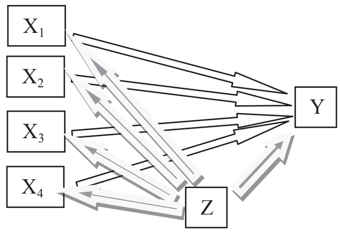

Figure 3.2 The common-cause fallacy in a multidimensional model

This result generalizes into other causal claims based on data-driven econometrics. The case studies discussed above show that today econometricians hardly ever test for *prima facie* Granger causality in a simple bivariate model. On the contrary, most studies aim at controlling for the influence of a few relevant factors. Using additional variables in model specification makes it immune, to some degree, from the fallacy, but the possibility that inferred causal claims are spurious still exists. For example, establishing the presence of probabilistic causality in a five-dimensional model (Figure 3.2) does not warrant that the causal conclusion holds in the case of a six-dimensional model including variable Z that may happen to be causally related to both variables X_1, X_2, X_3, X_4, and their purported effect, Y (cf. Sala-I-Martin 1997).[21] Therefore, the spuriosity of a causal claim based on probabilistic causality cannot in principle be excluded without additional evidence. This conclusion has substantial implications for policymaking.

Let me return to the case of testing *prima facie* causal claims. Economic policymaking is not targeted at influencing a bivariate process, but happens in a highly dimensional process underlying the actual economy. Given that econometrics establishes causal claims within an n-dimensional model, they can be false in an $n + m$–dimensional process that is targeted by economic policy.

Unfortunately, generalizing into the higher-order processes is not warranted. The causal claims that are true in a model of the n-dimensional process can be false in the $n + m$–dimensional world. To put it differently, the actual causal structure can differ from what is implied by causal claims inferred from data-driven econometric models. What follows is that this type of evidence does not warrant causal claims to be invariant under intervention, and hence policymaking requiring interventions changing the values of variables being relata of causal claims may lead astray. Unfortunately for the effectiveness of economic policymaking, models resulting from the common-cause fallacies have been used as evidence for intervening.

Today, the macroeconomic policy seems to rely on much more advanced modeling practices such as DSGE models (more on them in Section 5.2.3). However, 50 years ago, the situation was different. By then, fiscal and monetary policy had been based on a simple econometric model known as St. Louis equation specified as follows (Gordon et al. 1980):

$$Y_t = c + \sum_{n=0}^{2} M_{t-n} + \sum_{n=0}^{2} G_{t-n} + \varepsilon_t$$

Where:
M_{t-n} = n-lagged values of money
G_{t-n} = n-lagged values of government spending
c = constant
ε_t = error term

The estimation of this equation delivers response functions for the changes in government spending and money. Policymakers used these estimates to choose the amount of government spending (G) and taxes, and raise or lower monetary base so that the economy grows smoothly. What is easy to observe is that the St. Louis equation is a regression based on only two right-hand-side time series, and therefore resembles the case study of the reduced-form VAR model studying the dynamics of variables describing the number of casualties in the Israeli-Palestinian conflict. Let me discuss this out-of-fashion evidence for fiscal policymaking with a view to understanding how its misuse leads to unsuccessful interventions.

The St. Louis equation has been used for anticyclical policymaking, even though Sims (1972, 1992) has later shown that the causal structure presupposed by this regression results from the following spurious correlation: the VAR model including also inflation I and prices P does not deliver evidence for a statistically significant relationship between money M and output Y. Concluding that all the interventions based on the mis-specified regression have been unsuccessful is impossible due to the complexity of the economy. Even if a policy intervention is not transmitted by the purportedly causal relation that is its target, the expected result may occur because other factors may independently influence phenomena under consideration.

How to save yourself from interventions that fail? I have argued that econometric models presupposing a version of the probabilistic causality that is located between the *prima facie* and fully fledged concepts are likely, given the structure of the economic world (multicollinearity), to deliver causal claims resulting from spurious correlations. What follows is that these causal claims do not warrant intervention neutrality. In other words, interventions on money M to influence output Y will fail. However, it is not an argument for having probabilistic evidence in disregard.

Let me consider the following causal structure:

$$Y = F(G, I, P)$$

$$M = F(G, I, P)$$

If this structure remains constant, knowing (spurious) relation between M and Y can be successfully exploited. If the spurious correlation is observed between lagged values (e.g., $M_t \rightarrow M_{t+1}$), then such evidence can be fruitfully utilized.

Let me return to the case study of reduced-form VAR model of the dynamics between the number of Israeli and Palestinians killed in that conflict (Jaeger and Paserman 2008). Given the possibility of spurious causality, a policymaker is not entitled to intervene on right-hand-side variables to influence the left-hand-side variables. However, imagine that you manage a hospital located within the conflict zone. If the VAR model allows you to predict the number of casualties one day ahead, you are able to schedule the number of surgery teams.[22] Such policymaking does not act on the actual causal structure, but uses the incomplete knowledge of it with benefit for the policymaker. As long as the underlying causal structure producing the partial correlation between the two variables does not change, this type of policymaking would be useful even if it were based on a spurious causal claim.

3.4 Policymaking based on limited knowledge of causal structure

In this chapter, I have argued that causal claims based on atheoretical econometrics presuppose a version of the probabilistic definition of causality that locates itself between the *prima facie* and fully fledged conception requiring from causes to raise the probability of their effects, given all relevant factors. The atheoretical econometrics can be divided into the following two types of evidence. On the one hand, econometricians establish causal claims on the grounds of time series models that rely on calculating the partial correlations among the lagged values of variables. On the other hand, they estimate cross-sectional models that do not import causal structure from economic theory, but use the knowledge of time precedence of phenomena represented by variables and other extra-economic sources of evidence.

I have argued that policymakers in principle cannot exclude the possibility that causal claims based on the probabilistic evidence result from the common-cause fallacy and therefore are not invariant under interventions (when the system is kicked out from its usual state). Hence, such evidence can only be employed in support of policymaking actions that do not interfere with the causal structure producing phenomena of interest. My argument is of inductive nature. Lütkepohl's (1982) deductive result establishes that the presence of Granger causality in a two-dimensional model does not generalize into a three-dimensional model, including another causally relevant variable. Sims' (1972, 1992) empirical results suggest that the change from a three-dimensional model into a five-dimensional model can result in detecting that the three-dimensional model excluding other relevant variables establishes a spurious causal relationship. Given that any economic policymaking takes place in the economy rather than in an experimental setting, excluding factors not present in the econometric model being evidence for that claim, the policymaking intervenes on a more-dimensional process than the one modeled econometrically.

However, this argument does not support getting rid of probabilistic evidence from policymaking at all. Instead of the translating of causal claims from the probabilistic into the manipulationist meaning of causality and intervening on one variable (purported cause) to influence its effect, policymakers can employ such evidence to undertake the actions that do not intervene on the causal structure underlying the claim.

Notes

1 $P(E \mid C)$ – probability of *E*, given *C*; i.e., probability that even *E* will happen if *C* occurred. Some formulations modify the definiens as follows: $P(E \mid C) > P(E \mid \sim \mathrm{C})$ where $\sim \mathrm{C}$ denotes the probability that *C* does not happen.

2 In the sample I have studied (articles published in three top economic journals; i.e., *American Economic Review*, *Journal of Political Economics*, and *Quarterly Journal of Economics* from 2005–2015), there is no single case of the techniques developed in the DAG framework.

3 Many econometricians share the view that spurious (contrary to causal) correlations disappear when sufficiently long period of time is under analysis.

4 Considering that this section focuses on studying philosophical interpretation of research methods used in economics, the reader without a background in economics may wish to jump to the following Section 3.2 that studies these research methods in more detail. I have considered some related issues in Chapter 2, where I also distinguished between econometric techniques presupposing a version of the regularity view and probabilistic theories of causality. Here, I deal with problems arising from practicing atheoretical econometrics.

5 In support of this claim, I can cite how the authors consider the problem of common-cause fallacy (pp. 1598–1599). The lack of theoretical consideration and focusing on the features of data indicates the use of data-driven approach.

6 The authors have also calculated response functions of Israeli and Palestinians (section III.B). However, it only serves the role of supportive evidence.

7 The 'reduced-form' label is often omitted in the data mining literature where the VAR models are used for time-series analysis in an atheoretical way (Wei 2006) without their structural counterparts.

8 The analyzed case study exemplifies this widespread approach (Jaeger and Paserman 2008, p. 1594).

9 In detail, the authors tested for Granger causality for the following values of lags: 4,4; 7,4; 14,4; 21,4; 7,7; 14,7; 21,7; 14,14; 21,14; 21,21

10 The econometricians refer to the testing procedure as a Granger test, but the mention of Granger's 1969 paper (pp. 1590; 1594; 1598) indicates that they have in mind the testing procedure known as a Granger direct test.

11 In econometrics and statistics, the null hypothesis describes the situation whereby the estimates do not differ from 0 in a statistically significant way; i.e., the differences may result from random errors with the probability higher than presupposed by the researcher (usually 0.05 or 0.01).

12 The events described by variables repeat themselves in the sample, and therefore, the statistical procedure allow for estimating the probabilities of the occurrence of effects (Israeli retaliation attacks) for given causes (the number of Israeli fatalities).

13 The authors admitted that there are a few studies (e.g., Bowen and Rudenstine 1992) that deliver fragmentary knowledge on the determinants of dropout among doctoral candidates. There are also statistical comparisons of dropout rates between different disciplines accessible, but, at the time of publication, there were no studies focusing directly on the causes of resigning from Ph.D. programs.

14 Unfortunately, the authors remain silent on the approach to choosing their sample.

15 Considering that the only method of inference (apart from descriptive statistics summarizing the results of questionnaire offered in Table 4 and Table 5 [p. 464]) is this regression, the conclusions of the paper obviously refer to the probit regression.

16 The choice of the estimation procedure remains unspecified by Stock et al. (2006).

17 Partial correlation is a measure of a linear association (correlation) between two variables when the interrelation of other variables is controlled for. Therefore, discussing partial correlations is legitimate only in the case of linear econometric models.

18 Of course, the number of variables usually needs to be further reduced to refrain from overfitting the model to given data that limits the model's out-of-sample performance.

19 The definition states that C is a cause of E if and only if $P(E|CK) > P(E|K)$; i.e., considers all other factors (K) influencing an effect under consideration.

20 Given this, the criticism of causal claims based on a version of the probabilistic view only applies to the definition presupposed in econometric research and does not apply to the fully fledged definitions that are the topic of most philosophy of causality debates.

21 Testing whether a model is accurately specified reduces the chance that it represents a spurious correlation, but does not exclude such a possibility since some unobservable (immeasurable). Moreover, economists rarely test for the assumptions of statistical techniques they use.

22 Actually, such a use of this evidence requires further knowledge and decisions. For example, one would have to decide on the level of risk taking, given the funding obtained by the hospital.

References

Almond, D., Chay, K. Y., & Lee, D. S. (2005). The costs of low birth weight. *The Quarterly Journal of Economics*, *120*(3), 1031–1083. DOI: 10.1093/qje/120.3.1031

Atukeren, E. (2008). Christmas cards, Easter bunnies, and Granger-causality. *Quality & Quantity*, *42*(6), 835.

Baumgartner, M. (2008). The causal chain problem. *Erkenntnis*, *69*(2), 201–226. DOI: 10.1007/s10670-008-9113-2

Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinominal Models*. London: SAGE Publications.

Bowen, W. G., & Rudenstine, N. L. (1992). *In Pursuit of the Ph.D.* Princeton: Princeton University Press.

Canova, F. (1999). Vector autoregressive models: Specification, estimation, inference, and forecasting. In: Pesaran, H. & Schmidt, P. (eds.) *Handbook of Applied Econometrics Volume 1: Macroeconomics* (pp. 53–110). Oxford: Blackwell Publishers.

Cartwright, N. (1979). Causal laws and effective strategies. In: *Nous*, 419–437. DOI: 10.2307/2215337

Cartwright, N. (2001). What is wrong with Bayes nets? *The Monist*, *84*(2), 242–264.

Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.

Chang, T., Lee, C. C., & Chang, C. H. (2014). Does insurance activity promote economic growth? Further evidence based on bootstrap panel Granger causality test. *The European Journal of Finance*, *20*(12), 1187–1210. DOI: 10.1080/1351847X.2012.757555

Cramer, J. G. (1986). The transactional interpretation of quantum mechanics. *Reviews of Modern Physics*, *58*(3), 647. DOI: 10.1103/RevModPhys.58.647

Dawid, A. P. (2010). Beware of the DAG! *Proceedings of Workshop on Causality: Objectives and Assessment at NIPS 2008*, Whistler: Canada, PMLR, 6, pp. 59–86.

Demiralp, S., & Hoover, K. D. (2003). Searching for the causal structure of a vector autoregression. *Oxford Bulletin of Economics and Statistics*, *65*, 745–767. DOI: 10.1046/j.0305-9049.2003.00087.x

Dupré, J. (1984). Probabilistic causality emancipated. *Midwest Studies in Philosophy*, *9*(1), 169–175.

Eells, E. (1991). *Probabilistic Causality*, Vol. 1. Cambridge: Cambridge University Press.

Ehring, D. (2009). Causal relata In: Beebee, H. et al. (eds.) *Oxford Handbook of Causation* (pp. 387–413). Oxford: Oxford University Press.

Engle, R., Hendry, D.,& Richard, J.-F. (1983). Exogeneity. *Econometrica*, 277–304. DOI: 10.2307/1911990

Good, I. J. (1959). A theory of causality. *British Journal for the Philosophy of Science*, *9*, 307–310.

Good, I. J. (1961a). A causal calculus (I). *The British Journal for the Philosophy of Science*, *11*(44), 305–318.

Good, I. J. (1961b). A causal calculus (II). *The British Journal for the Philosophy of Science*, *12*(45), 43–51.

Gordon, R. J., Okun, A. M., & Stein, H. (1980). Postwar macroeconomics: The evolution of events and ideas. In: *The American Economy in Transition* (pp. 101–182). Chicago: University of Chicago Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424–438. DOI: 10.2307/1912791

Granger, C. W. J. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamic and Control*, *2*(4), 329–352. DOI: 10.1016/0165-1889(80)90069-X

Granger, C. W. J. (1988). Some recent development in a concept of causality. *Journal of Econometrics*, *39*(1–2), 199–211. DOI: 10.1016/0304-4076(88)90045-0

Granger, C. W. J. (2012). Forecasting. In: Mäki, U. (ed.) *Philosophy of Economics* (pp. 311–327). London: Elsevier.

Hardt, Ł (2017). *Economics without Laws*. Cham: Palgrave Macmillan.

Hausman, D. (1998). *Causal Asymmetries*. Cambridge: Cambridge University Press.

Henschen, T. (2018). The in-principle inconclusiveness of causal evidence in macroeconomics. *European Journal for Philosophy of Science*, *8*(3), 709–733.

Hesslow, G. (1976). Two notes on the probabilistic approach to causality. *Philosophy of Science*, *43*(2), 290–292. DOI: 10.1086/288684

Hicks, J. (1979). *Causality in Economics*. New York: Basic Books.

Hoover, K. D. (2001). *Causality in Macroeconomics*. Cambridge: Cambridge University Press.

Hoover, K. D. (2003). Nonstationary time series, cointegration, and the principle of the common cause. *The British Journal for the Philosophy of Science*, *54*(4), 527–551. DOI: 10.1093/bjps/54.4.527

Hoover, K. D. (2008). Causality in economics and econometrics. In: Durlauf, S. & Blume, L. (eds.) *The New Palgrave Dictionary of Economics*. New York: Palgrave Macmillan.

Jaeger, D. A., & Paserman, M. D. (2008). The cycle of violence? An empirical analysis of fatalities in the Palestinian-Israeli conflict. *American Economic Review*, *98*(4), 1591–1604. DOI: 10.1257/aer.98.4.1591

Keynes, J. M. (2013). *A Treatise on Probability*. New York: Dover Publications, Inc.

Leamer, E. E. (1985). Vector autoregressions for causal inference? *Carnegie-Rochester Conference Series on Public Policy*, *22*, 255–304. North-Holland.

Lee, H., Lin, K.,& Wu, J. (2002). Pitfalls in using Granger causality tests to find an engine of growth. *Applied Economics Letters*, *9*(6), 411–414. DOI: 10.1080/13504850110088132

Lefgren, L., Sims, D., & Lindquist, M. J. (2012). Rich dad, smart dad: Decomposing the intergenerational transmission of income. *Journal of Political Economy*, *120*(2), 268–303.

Lütkepohl, H. (1982). Non-causality due to omitted variables. *Journal of Econometrics*, *19*(2–3), 367–378. DOI: 10.1086/666590

Lütkepohl, H. (1990). Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. *The Review of Economics and Statistics*, 116–125. DOI: 10.2307/2109746

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Maddala, G. S., & Lahiri, K. (2009). *Introduction to Econometrics*. Hoboken, NJ: Wiley.

Maziarz, M. (2015). A review of the Granger-causality fallacy. *The Journal of Philosophical Economics*, *8*(2), 86–105.

Maziarz, M. (2018). Causal inferences in the contemporary economics. *Mendeley Data.* Retrieved from: http://dx.doi.org/10.17632/v7dhjnd8xg.2. Access: 16th October 2018.

Maziarz, M. (2019). It's all in the eye of beholder. *Argumenta Oeconomica*, *2*(43), 307–328. DOI: 10.15611/aoe.2019.2.13

McCrorie, R.,& Chambers, M. (2006). Granger causality and the sampling of economic processes. *Journal of Econometrics*, *132*, 311–336. DOI: 10.1016/j.jeconom.2005.02.002

Moneta, A. (2005). Causality in macroeconometrics: Some considerations about reductionism and realism. *Journal of Economic Methodology*, *12*(3), 433–453. DOI: 10.1080/13501780500223742

Nowak, S. (1960). Some problems of causal interpretation of statistical relationships. *Philosophy of Science*, *27*(1), 23–38. DOI: 10.1086/287710

Pearl, J. (1982). *Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach.* Cognitive Systems Laboratory, School of Engineering and Applied Science, University of California, Los Angeles, pp. 133–136.

Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press.

Pearl, J., & Verma, T. S. (1995). A theory of inferred causation. *Studies in Logic and the Foundations of Mathematics*, *134*, 789–811. DOI: 10.1016/S0049-237X(06)80074-1

Reichenbach, H. (2012a). *The Direction of Time.* Berkley: University of California Press.

Reichenbach, H. (2012b). *Hans Reichenbach: Selected Writings 1909–1953*, Vol. 2. Dordrecht: Reidel Publishing Company.

Sala-i-Martin, X. X. (1997). I just ran four million regressions Working Paper No. w6252. Washington, DC: National Bureau of Economic Research. DOI: 10.3386/w6252

Reiss, J. (2007). *Time Series, Nonsense Correlations and the Principle of the Common Cause* In: Russo, F. & Williamson, J. (eds.) Causality and Probability in the Sciences. London: College Publications, pp. 179–196.

Salmon, W. C. (1980, January). Causality: Production and propagation. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1980*(2), 49–69. DOI: 10.1086/psaprocbienmeetp.1980.2.192586

Sala-i-Martin, X. X. (1997). I just ran four million regressions Working Paper No. w6252. Washington, DC: National Bureau of Economic Research. DOI: 10.3386/w6252

Schechter, L. (2007). Theft, gift-giving, and trustworthiness: Honesty is its own reward in rural Paraguay. *American Economic Review*, *97*(5), 1560–1582. DOI: 10.1257/aer.97.5.1560

Schelling, Th. (1960). *The Strategy of Conflict.* London: Oxford University Press.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 238–241. DOI: 10.1111/j.2517-6161.1951.tb00088.x

Sims, Ch. (1972). Money, income and causality. *The American Economic Review*, *62*(4), 540–552.

Sims, Ch. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48. DOI: 10.2307/1912017

Sims, Ch. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, *36*(5), 975–1000. DOI: 0.1016/0014-2921(92)90041-T

Skyrms, B. (1980). *Causal Necessity.* New Haven and London: Yale University Press.

Sober, E. (2001). Venetian sea levels, British bread prices, and the principle of the common cause. *The British Journal for the Philosophy of Science*, *52*(2), 331–346. DOI: 10.1093/bjps/52.2.331

Spirtes, P., Glymour, C., & Scheines, R. (2012). *Causation, Prediction, and Search.* New York: Springer-Verlag.

Stahlman, W. D. (1948). *Wiener's Theory of Prediction for Discrete Time Series.* Doctoral Dissertation, Massachusetts Institute of Technology, Department of Mathematics, Cambridge, MA.

Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *The Journal of Economic Perspectives*, *15*(4), 101–115. DOI: 10.1257/jep.15.4.101

Stock, W. A., Finegan, T. A., & Siegfried, J. J. (2006). Attrition in economics Ph. D. programs. *American Economic Review*, *96*(2), 458–466. DOI: 10.1257/000282806777212044

Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Publishing Company.

Verbeek, M. (2012). *A Guide to Modern Econometrics*. Rotterdam: John Wiley and Sons.

Wei, W. W. (2006). Time series analysis. In: Little, T. (ed.) *The Oxford Handbook of Quantitative Methods in Psychology*. Oxford: Oxford University Press.

Wiener, N. (1956). The theory of prediction. In: Beckenbach, E. (ed.) *Modern Mathematics for the Engineering* (pp. 165–190). New York: McGrew Hill.

Williamson, J. (2009). Probabilistic theories of causality. In: Beebee, H. et al. (eds.) *The Oxford Handbook of Causation* (pp. 185–212). Oxford: Oxford University Press.

Zellner, A. (1988). Causality and causal laws in economics. *Journal of Econometrics*, *39*(1–2), 7–21. DOI: 10.1016/0304-4076(88)90038-3

4 Counterfactuals

Similar to the regularity view, the counterfactual approach to causality was established by David Hume (1992, p. 51), who, after coining the regularity-view definition of causality, added: "[o]r in other words, where, if the first object had not been, the second never had existed." In other words, the second Humean definition defines causal relations in the following way: *A* causes *B* if and only if the following counterfactual statement is correct: if *A* had not existed, *B* would not happen. Hume's two definitions following each other produced much confusion among philosophers of causality because they are not synonymical, despite the connection "in other words" employed by Hume. Possibly, he voiced the two definitions with the aim of distinguishing epistemic and ontological views. According to such an interpretation, the regularity view is an epistemological stance: causality can be inferred by means of observing regular constant conjunctions where causes precede their effects in time. According to the latter definition, the nature of causality is not the time precedence, but the fact that a cause produces its effect so that without its cause, the effect does not happen. However, Galen Strawson (1987, p. 193) convincingly argued that, for Hume, these two definitions of causality are limited to indicating observable appearances of causal relations. Aside from Hume's view, the counterfactual view is often considered as an ontological position. For example, Laurine Paul (2009, p. 158) acknowledged that the counterfactual analysis "captures something essential and fundamental about . . . causation."

The chapter is structured as follows. Section 4.1 focuses on reviewing the philosophical literature devoted to studying causal relations in terms of counterfactuals, and discusses its applications to causal inference, and the influence of the counterfactual approach on the philosophy of economics. Section 4.2 discusses two case studies of economic research that produce causal claims understood in line with the counterfactual approach. These are (1) establishing a counterfactual claim on the grounds of a previously calibrated theoretical model, and (2) employing a case study approach. Section 4.3 discusses policy-oriented implications on the basis of previous analyses.

4.1 Counterfactual conditionals and causality

In this section, I analyze the philosophical literature developing Hume's second definition of causality and studying how can we deliver evidence in favor of a counterfactual claim. First, I review the main philosophical theories of counterfactual causality (Section 4.1.1) and discuss a few significant problems with the counterfactual approach and responses to these problems (4.1.2). Later, I review the discussions from the philosophy of economics literature that are relevant to the counterfactual view on science. The following subsections exemplify the inference of counterfactual causal claims and discuss their use for policy.

4.1.1 Classical formulation

In contemporary philosophy, David Lewis (1973) popularized the counterfactual approach to causality. Lewis acknowledged Hume's second definition but refused to consider it as an ontological definition of causality: "I take Hume's second definition as my definition not of causation itself, but of causal dependence among actual events" (p. 563). For Lewis, the counterfactual condition delivers epistemic guidance in causal search. According to the counterfactual definition of causality, '*C* causes *E*' is correct if the following statement is true: if *E* did not happen, then *C* would not happen. The counterfactual condition is usually interpreted in the following way: if event *C* were miraculously erased from the world and it would be the only change (everything else stays the same), then *E* would not happen if and only if *C* causes *E* (cf. Reiss 2015, p. 12). What is crucial for further discussion is that such a formulation requires causes to be necessary conditions of their effects (cf. Chapter 2). Lewis (1973), instead of thought-experimenting with a miraculous intervention, employed modal logic and considered two possible worlds w_1 and w_2. If the two worlds are alike with the exception that event *C* did not happen in w_1 but happened in w_2 and that the event *C* was followed by *E* in w_2 and did not happen in w_1, then *C* causes *E* (cf. Morgan and Winship 2014, p. 42). In other words, in order to decide whether *C* caused *E*, one should consider what would happen in the closest parallel universe in which *C* did not happen.

The problem with such an interpretation is that it is 'metaphysical' in the logical-positivist sense; i.e., taking counterfactuals as descriptions of parallel universes makes their truth-value impossible to identify. Therefore, it seems that the Humean stance (taking the regularity definition as an epistemic concept and the counterfactual definition as an ontological position) seems appropriate. Furthermore, Daniel Hausman (1998, p. 112) highlighted that "[t]he 'comparative overall similarity' of possible worlds is a vague relation." Even if such a comparison were possible, the relation of similarity is subjective. Similar to the problem of defining 'class' for analyzing the relation of regularity between classes of events, the overall comparative similarity is an ambiguous term.

Lewis' earlier account was also criticized for not incorporating an account of time asymmetry. Many philosophers take the unidirectional nature of causal relations as their characteristic feature (cf. Frisch 2005, p. 21). In a later study. Lewis (1979) argued in favor of time asymmetry of causality by considering the asymmetry of overdetermination and phenomena depicted by wave theory and stated that "[s]eldom, if ever, can we find a clearly true counterfactual about how the past would be different if the present were somehow different" (p. 445).

John Mackie, the philosopher who coined the INUS condition (1965) described in Chapter 2, put forth later a counterfactual account similar to Lewis' views. He defined a cause as "what makes the difference in relation to some assumed background or causal field" (Mackie 1974, p. 71). Mackie offered the conception to coin an account capable of analyses of singular causation. Despite being more fruitful in depicting the dependence of counterfactual conditions on background knowledge, it is Lewis' approach that gained popularity (Salmon 2006, p. 130). However, such simple accounts of causality as the early version of Lewis' views or Mackie's analysis faced considerable criticism.

A widely known problem (the problem of overdetermination) with the counterfactual approach to causality is a thought experiment of two people throwing a rock into a glass window. Since none of the two events fulfill the counterfactual condition (the window would break if throwing a rock by the first or second person did not happen), then accepting this definition of causality leads to the conclusion that rock-throwing is not a cause of breaking a window. Another counterexample is the case of preemption. Consider that the end of the recession was caused by lowering interest rates, but, if not the monetary policy, the free-market activity would cause the recovery. In such a situation, monetary policy cannot be considered a cause because it does not fulfill the counterfactual condition. Some of the consequences of the counterfactual account are counterintuitive. For example, analyzing cases of causation by omission (depicted as 'if *C* happened, then *E* would not) leads to the conclusion that causality can be possible without a causal relation since relations can exist only among its relata (Lewis 2004). The limited ability to explain phenomena by this approach is highlighted by those who accuse it of being circular (e.g., Woodward 2005). As Paul (2009, p. 172) put it, "one's account of the causal relation in terms of counterfactual dependence requires an account of counterfactual dependence in terms of causation."

4.1.2 Recent developments

Due to these and other obstacles, Lewis (2000) revised his earlier views. Contrary to his earlier account based on considering if events *C* and *E* occur, the new approach is wealthier in also employing the issue of space and time (the 'where' and 'when' questions). Furthermore, Lewis' new theory of causation overlaps with the mechanistic approach discussed in the next chapter. Lewis (2000) defined his new theory of causality as follows:

> [w]here *C* and *E* are distinct actual events, let us say that *C* influences *E* if and only if there is a substantial range C_1, C_2 . . . of different not-too-distant

> alterations of C (including the actual alteration of C) and there is a range E_1, E_2 . . . of alterations of E, at least some of which differ, such that if C_1 had occurred, E_1 would have occurred, and if C_2 had occurred, E_2 would have occurred, and so on. Thus we have a pattern of dependence of how, when, and whether upon how, when, and whether.
>
> (p. 190)

In addition to including the requirement of spatiotemporal connection, Lewis (2000) developed his counterfactual account of causality by including the 'step-wise influence,' which makes this counterfactual theory resemble the process or mechanistic theories (cf. Chapter 5). Reconsidering the thought experiment of two people throwing a rock at a glass window leads to the conclusion that only the rock-throwing that caused the window to shatter is the cause. Peter Menzies (2017) indicated that the new theory is better at handling some of the counterexamples voiced against Lewis' earlier account.

Most of the counterfactual theories analyze deterministic causal relations, but there are also stochastic theories (Paul 2009). For instance, Morgan and Winship (2014, p. 380 et seq.) considered a simple two-equation linear model and argued that producing out-of-sample prognoses is grounded in the counterfactuals of the following form: if X were x, then Y would, on average, be y. However, they do not voice their opinion on how the counterfactual account relates to the regularity approach. Should the counterfactual causal claim be valid on average? Alternatively, should they include the *ceteris paribus* clause? These topics wait to be addressed, given that the counterfactual definition of causality is usually employed for justification of singular causal statements. The following subsection focuses on the reception of the counterfactual causation by economists and the methods of causal inferences grounded in this stance.

4.1.3 Counterfactuals in the philosophy of economics

In the contemporary philosophy of economics, the emerging consensus states that counterfactual causal claims are employed for describing token-level dependencies or singular causation; i.e., the possible or actual situations when an instance of one phenomenon or event causes another. The use of singular causal claims in economics ranges from evaluating a given economic policy, indicating a cause of an event, to discussing the outcomes of an intervention. An example of such a statement is a claim that the recent financial crisis caused the 2007–2008 economic slowdown. For instance, Julian Reiss (2015, p. 11) argued that:

> throughout the social and biomedical sciences, researchers focus on singular causal analysis at least some of the time. . . . [E]conomists are interested in the causes of business cycles but also why a specific event such as the Great Recession occurred; case-study research, often using only a single case, is ubiquitous in the social sciences. . . . [W]e need an account of singular causation in addition to an account of causal laws.

Nancy Cartwright (2007) voiced a similar opinion that "[c]ounterfactuals are a hot topic in economics today" (p. 191) and are used to address the question 'what if?' such as: "what if the policy were put in place?" (p. 193).

I take issue with the view that counterfactual causal claims are of significant importance in economics. Economists[1] are usually interested in general, type-level causality (e.g., Henschen 2018) rather than singular, token-level relations. Counterfactuals play a role in formulating conclusions about singular instances. The interest in singular causation is usually present in the activities of economists not related to causal inference, but rather using causal knowledge to some purposes such as delivering policy advice or predicting. Counterfactual causality plays a considerably limited role in research at the frontier of economics. However, the reliance of other sciences on counterfactual causality is stronger. For example, epidemiology utilizes the potential outcome approach to delivering evidence in favor of general causal claims (Little and Rubin 2000). The potential outcome framework defines causes as effects of manipulations (exposures in the epidemiologist dictionary) and, to some degree, overlaps with James Woodward's (2005) manipulationist theory of causality (cf. Chapter 6) that merges counterfactual and manipulationist approaches.

Even though the philosophy of epidemiology falls beyond the scope of this book, discussing the potential outcome approach is useful for a better understanding of the method of causal inference employed in the first case study. Broadbent (2013, p. 1) defined epidemiology as "the study of the distribution and determinants of disease and other health states in human population using group comparisons for the purpose of improving population health." The feature of epidemiology that it shares with econometrics; i.e., studying observational data, or at least data produced by interventions located beyond the control of researchers, is not implicit in this definition. The potential outcome approach was coined initially to analyze the results of experimental studies by Jerzy Neyman and subsequently generalized by Donald Rubin (1990) into the context of observational research (a.k.a. the Rubin causal model). Roughly speaking, statistical techniques belonging to the potential outcome approach aim to minimize the effect of nonrandom assignment to treatment and control groups.

Cartwright (2007) distinguished between Galilean counterfactuals and the implementation-specific ones. The distinction proves fruitful in discussing counterfactual causal claims in economics and elsewhere. The 'Galilean' counterfactuals[2] address a theoretical question of what the state (e.g., value) would be of an effect (variable) if its cause were in a different state (value). The Galilean counterfactual claims are not suitable for discussing the effects of actual interventions. This role is played by the manipulationist counterfactuals that aim explicitly to depict the results of actual interventions and address the question of what would happen if an intervention were conducted. In other words, the former type of counterfactual claims describes what would happen if an aspect of the world were different while the latter predicts the results of intervention (what would happen if X were set to x). The following subsection focuses on discussing the two main types of evidence for counterfactual claims found in economics. I argue that the claims are of Galilean nature.

4.1.4 Philosophical views on inferring counterfactuals

Like any other concept of causality, the versions of the counterfactual approach can either serve as an epistemic view serving the purpose of guiding research or as an ontological description of what causality is. As long as the latter interpretation of the counterfactual theories locates itself beyond the scope of the discussion, this section focuses on the previous topic. There are two main approaches to producing evidence for a counterfactual causal claim. First, one can support such a claim with a previously established causal model. Second, counterfactual claims can be established on the grounds of a qualitative study. As we will see, both approaches raise some difficulties. The use of Lewis' concept of similarity between parallel universes is clearly of limited help. How can one know that without C, E would not happen?

Supporting counterfactuals with models

The philosophers and methodologists concerned with the problem of how to produce counterfactuals on the grounds of previously obtained knowledge usually focus on DAGs and structural-equation models. While these models are rarely used in contemporary mainstream economics, some conclusions also apply to make counterfactual causal claims from either econometric or axiomatic (theoretical) models. For instance, Ilya Shpitser and Judea Pearl (2012) strived for formalizing testing counterfactual conditions with the use of DAG. In general, if this method of inferring counterfactual statements is employed, the Galilean counterfactuals are justified to a degree to which the underlying DAG model is believed to be accurate. However, the case of the policy counterfactuals of the form 'what would happen if policy $X = x$ were introduced?' is epistemically more demanding. There is nothing in the data that indicates if an identified causal relation $X \rightarrow Y$ is implementation neutral. At least some of policy interventions modify causal structure.

In order to establish causal counterfactuals, one is to assume that the underlying structural-equation model (causal model) is accurate and deduce a counterfactual claim under consideration from that model. As Cartwright (2007, p. 195) put it:

> [t]o evaluate counterfactuals of this kind we need a causal model; and the causal model must contain all the information relevant to the consequent about all the changes presumed in the antecedent. There is no other reasonable method on offer to assess counterfactuals.

In a similar vein, James Heckman (2000, 2008) highlighted that the interpretation of the structural econometric model conducted for economic policymaking is grounded in the counterfactual approach. Such analyses conclude with statements of the form 'if X were (were set to) x, then Y would be y' (cf. Morgan and Winship 2014, p. 49). This interpretation bears a resemblance to the manipulationist theory put forward by James Woodward (2005), cf. Chapter 6).

Pearl (2009) also employed the counterfactual account to analyze interventions in the structural-equation framework. The following distinction should be made here. On the one hand, the use of structural-equation framework (or – more widely – theory-inspired econometrics) to discover the structure of causal laws (general causation) governing an economic system is grounded in the regularity view (cf. Chapter 2). On the other hand, claims about singular (token-level) causal claims such as 'lowering interest rates at time t will accelerate economic development' are grounded in the counterfactual approach since the example mentioned earlier can be reformulated into the counterfactual condition: 'if interest rates were lowered, the economy would grow so quickly.' The contemporary methods of causal inference are, according to Cartwright (2007), too limited to deliver evidence for the laws of the latter kind.

In order to justify a singular causal statement, we need to have the right model depicting laws of the universe (nomological model, cf. Lewis 1983). Therefore, several philosophers claim that employing the counterfactual approach to causality does not enlighten the discussion of how to infer causal claims: singular causal claims are justified to the degree that laws covering a given causal claim are justified. For instance, Cartwright (2007, p. 196) argued that the counterfactual statements can be grounded in the structural equation models, and they are only true if the underlying causal model grasps the interdependencies rightly: "they [the structural models] are to be functionally correct and to provide a minimal full set of causes on the right-hand side for the quantity represented on the left."

King and Zeng (2007) considered the fallibility of counterfactual inferences. One of their conclusions that is of interest for this book states that the justification of predictions based on counterfactuals depends on how far the employed counterfactuals deviate from evidence (both models and data). For instance, the counterfactuals discussing how a 1% change of inflation will influence economic development are more justified than predicting the effects of a 100% change. Here, a useful distinction between the out-of-model risk and model risk (cf. Morini 2011) coined on the grounds of quantitative finance should be considered. The former type of risk denotes all the influences external to a model. The latter includes all the factors included. In general, the more a possible intervention exceeds a model (e.g., available data points used for estimation), the less likely a counterfactual statement is to turn out right. However, even a counterfactual describing a small deviation from a model may be false if the conditions external to the model change.

There is another problem with the latter type of counterfactual claims employed to justify economic policymaking. In order for such claims to be valid in case of conducting actual interventions instead of 'what-if' theorizing, causal/nomological models must depict laws that are invariant under intervention. As I have argued in Chapter 3, econometric models of observational data are unable to deliver decisive evidence for intervention-neutrality. There are three sources of potential error. First, the inferences from observational data are susceptible to the common-cause fallacy. Second, a system under study may behave differently under intervention than it would when it was left alone. Finally, a counterfactual condition can be implemented in several different ways. For example, a 1% reduction in inflation will have very different effects when it is achieved by monetary policy (e.g., by raising interest rates) or by a decree freezing prices at the last year level.

Cartwright (2007, p. 201) offered a solution according to which such a problem can be disentangled by putting forth 'implementation-specific' counterfactuals defined as an answer to the exact question 'what would happen if policy *X* were introduced?'. How could such a counterfactual be assessed? Alternatively, what type of evidence is required to justify them? Experimentation – or, at least, observing the results of similar interventions – seems to be the best way. Nancy Cartwright and Julian Reiss (2008) developed the criticism of three popular accounts of counterfactual causality. For example, they criticized Pearl's (2000) counterfactual semantics developed within the DAG framework on the grounds that discussing policy changes as changes in a single variable is rarely possible in economics, especially if expectations are present within a model.

Supporting singular causal claims

Apart from drawing singular (token-level) counterfactual claims on the grounds of general (type-level) causal models, such claims can be inferred from qualitative studies of single cases. Gary Goertz and Jack Levy (2007) argued that these two types of causal inference (here classified as utilizing a version of the counterfactual view) presuppose two different views of causality. They located the demarcation line between the statistical sciences grounded in the probabilistic or regularity views and case-study analysis that presuppose the cause and effect relationship in terms of the counterfactual condition: "there are two basic schools of thought on causation that are relevant. One is the covering law, statistical/probabilistic causation school. The second is the necessary condition, counterfactual approach" (Goertz and Levy 2007, p. 18). The problem with qualitative research seems to lie in the indeterminacy of necessary conditions: there are always many potential causal factors and, at least in principle, choosing the relevant ones may be impossible based on case study analysis. Goertz and Levy (2007, p. 15) believe that the situations in the real-life case-study analyses are epistemically more accessible than in the case of abstract methodological considerations because "the goal is to focus on one important causal factor. The aim is not a 'complete' explanation of the event but rather a more modest one of exploring the consequences of a key independent variable." Therefore, such a reasoning, assuming that counterfactual causes are necessary conditions, can be reduced to justifying the claim that *A* is a necessary and nontrivial condition for *B*. Such reasoning is usually justified with either considering theoretical, aprioristic knowledge on connection between *A* and *B*, or on simple inductive methods such as comparing similar cases in the case of comparative case study (cf. Dul and Hak 2008, p. 139).

Mainstream economists do not employ the method of case-study analysis often. A possible reason is a fact that such analyses are qualitative, 'soft,' holistic, and may be seen as unscientific due to being insufficiently mathematized (cf. McCloskey 1998). Nevertheless, case studies are widely used in the less quantitative social sciences such as management and the political sciences. A case study is "an intensive study of a single unit to generalize across a larger set of units" (Gerring 2004, p. 341). One of the features of this kind of analysis is being oriented on process tracing (p. 342). However, the discussion of the place of case-study analysis among different methods of causal inference and its role in the scientific endeavor is far from being settled. Bennett (2004, p. 22) distinguished between case studies employing process tracing and counterfactual analysis. The process-tracing case studies are aimed to assess whether a causal mechanism operates in a studied case in line with a theoretical description. In other words, their purpose is to test a theory. Despite the name suggesting explicit grounding in the process/mechanistic theories, process-tracing case studies overlap with Lewis' (2000) counterfactual theory of causality that epistemically defines causality as the counterfactual relation between *C* and *E*, adding the requirement of a chain of connections between *C* and *E* (cf. Bennett 2004, pp. 22–23).

Another issue that awaits to be addressed is the question of whether a qualitative study of a single case is capable of enlightening causal relations. John Gerring (2004, p. 346) indicated that only cross-unit case studies are capable of delivering causal inferences. On the contrary, single-case analyses can only deliver detailed knowledge. In his later book, Gerring (2006, p. 151) explained that case studies employed for causal inference overlap with quasi-experimental studies: they are a method of identifying difference-making factors. Bennett (2004, p. 25) opposed this opinion and distinguished between two types of case-study analysis grounded in the counterfactual approach to causality. Idiographic case-study counterfactuals are a method of qualitative research that produces causal inferences based on a detailed analysis of a single case. Such analyses employ qualitative inquiry that is aimed at theorizing a possible connection between *C* and *E* in order to establish a counterfactual statement. Nomothetic-counterfactual studies, on the contrary, are based on employing a widely accepted theory to produce singular-causal counterfactuals.

Gerring (2006, p. 213) defined causes in terms of necessary conditions which indicate that some formulations of the counterfactual approach overlap with the regularity view and, especially, defining causes as necessary or sufficient conditions. He indicated that case studies are helpful in causal inference only if an identified relation is deterministic. In such a case, a comparative case study can enlighten our knowledge of the difference-making factor. However, in order to put faith in a causal generalization grounded in a case-study analysis, one should possess a priori knowledge (or establish it by either statistical or theoretical means) about the determinism of the relationship. Interestingly, even some of those who use the method of case study in their research are skeptical about the possibility of causal inference (e.g., Dion 2003). Considering that case studies are sometimes interpreted as serving the purpose of theory-building (Bennett 2004) and keeping in mind other limitations of (especially policy-oriented) counterfactual statements, causal inferences based on case studies should be treated with a considerable dose of skepticism.

Similar to any other approach to causal inferences, employing the counterfactual definition faces considerable criticism. Alexander Philip Dawid (2000) argued that counterfactual claims are, by definition, impossible to test, and therefore their use can be misleading. He distinguished between two types of counterfactuals. The first type, which resembles Cartright's manipulationist counterfactual, can be labeled policy-oriented counterfactuals (e.g., does funding a bailout end financial crisis?). addressing the question of whether doing *X* will result in *Y*. The second type (assessing a policy; e.g., did the bailouts end the 2007–2008 financial crisis?) asks if *Y* was caused by *X* or by some other factor. In such a situation, the counterfactual claim is as follows: Would *Y* happen if *X* did not happen?'. The former type is an inference about the effects of causes (Dawid 2000, p. 408). The latter, on the contrary, focuses on the causes of effects.

In summary, counterfactuals are often used in economics to put forth singular causal statements. However, such statements need an underlying nomological model for their justification: one can rely on such counterfactuals as much as one believes that the underlying model is accurate. Possibly except for a few types of case-study analyses aimed at theorizing on a causal connection in line with Lewis' (2000) theory that is employed in other social sciences (Fearon 1991), inferring singular causal statements is impossible without having prior knowledge. Generally speaking, the two divergent ways of employing the methods of causal inference grounded in the counterfactual approach to causality are (1) producing 'what if?' conditionals based on already possessed knowledge, and (2) putting forth singular causal statements on the grounds of qualitative research. A prediction of a result of a macroeconomic intervention grounded in a structural equation model exemplifies the former approach. In this case, the truth of the counterfactuals depends on the background evidence: such statements are valid only if one possesses the right nomological model (Cartwright 2007). Therefore, this approach to producing counterfactual statements is, de facto, a method of inference from previously obtained evidence for type-level causal claims to singular (token-level) causal claims. For instance, Hausman (1998, p. 119) indicated that knowledge about counterfactuals can be employed to produce predictions. Unfortunately, the question regarding what evidence is needed to establish a contextualized policy counterfactual waits for an in-depth analysis. I will next elaborate on this view in detail based on two case studies. In the following section, I discuss the use of counterfactuals for policy.

4.2 Counterfactuals and causal inference in economics

Economists use counterfactuals to put forth causal claims that describe token-level relations (i.e., singular instances rather than types are relata). There are two approaches to delivering evidence for such claims. First, singular causal claims can deductively depend on (results from) a model depicting type-level causal relations.

Second, they can result from a detailed study of a particular instance. In this section, I instantiate these two main approaches to supporting singular causal claims with two types of causal claims. Atif Mian and Amir Sufi (2012) put forward a counterfactual causal claim on the grounds of their econometric model (Section 4.2.1) aimed at analyzing the influence of fiscal stimulus ('cash for clunkers') on economic development. The counterfactual analysis assessed the effects of the stimulus package by comparing car purchases in cities most and least affected by the program. Heidi Williams (2013) conjectured about the influence of intellectual property rights on the innovation of the human genome project, establishing the token-level causal claim on the grounds of a case study[3] (Section 4.2.2).

4.2.1 What would be the level of economic development if a stimulus were not introduced?

In the aftermath of the 2007–2008 financial crisis, governments were struggling with insufficient levels of demand and have undertaken several anti-recessionary measures. In the United States, federal administration in cooperation with Federal Reserve-stimulated growth with excessive bailouts for investment banks troubled with bad debt and programs aimed at stimulating the consumption of durable goods. An example of the latter is Cars Allowance Rebate System (CARS, usually known as the 'cash for clunkers' program) that relied on payments[4] for car dealers for older cars traded in by consumers. Atif Mian and Amir Sufi's (2012) study is an attempt at assessing the effects of this stimulus program on the number of purchased cars. The counterfactual language enters their paper from the very beginning, where its purpose is specified as addressing the question: "[w]hat would have been the trajectory of economic variables in the absence of the stimulus program?"

Unfortunately, assessing such a question is not an easy task. The opinions of theoretical economists on the effectiveness of such stimuli packages are divided (Blinder 2008; Becker 2010). Empirical assessments are also harsh and potentially misleading, because the cash for clunkers program was at work only for a few months of the summer of 2009. Therefore, the possibility that other factors (e.g., interest-free financing offered by a few car manufacturers at the time (Bunkley 2010) influenced the growth of car sales observed in that period. Furthermore, the short period when the program operated makes it possible that the sales observed in the period were not additional, but rather 'borrowed' from the following months. If this explanation is valid, those consumers who planned to buy a new car regardless of the program decided for the purchase earlier, but the number of additional purchases is marginal. Given these explanations for the possibility of a spurious rise of car sales resulting from the operation of other causes and divided opinions, a direct inference from time-series aimed at comparing the time when the cash for clunkers program was present to the sales recorded before and after the program may be misleading – considering that the program operated only in the United States, employing panel-data comparisons to other regions of sufficient similarity is also impossible.

Taking the limitations of other research designs to heart, Mian and Sufi (2012) designed an econometric study in a way that resembles experimental study (cf. Section 6.2). To stand as proxies for the experimental and control groups, Mian and Sufi (2012) used a self-constructed variable denoting the number of cars fulfilling the requirements of the program (i.e., clunkers sufficiently old and inefficient) and divided the regions into high- and low-clunker cities. This approach is based on the conjecture that if only a few 'clunkers' (i.e., cars liable for buy-out) are in a city, then the effects of the program will be limited in that city. The first step of the analysis was to estimate the number of cars in each administrative region. To do so, the authors used (1) the list of cars (and their production years delivered by edmunds.com) that fulfill the criteria and can be traded in at the time of purchasing a new car, and (2) the numbers and types of cars registered in each 'city'[5] from R.L. Polk. Interestingly, these two sources of data use different types of grouping observations,[6] so merging datasets required data preprocessing.

The econometric panel models estimated by Mian and Sufi (2012) aim to estimate the marginal effect of the CARS program using the following thought experiment:

> [s]uppose there are two cities, one in which everyone owns a clunker and one in which no one owns a clunker. The 'experiment' uses the nonclunker city as a control to assess the counterfactual level of auto purchases in the absence of the program for the city with clunkers.
>
> (p. 1118)

Their approach utilizes the requirements of the program: only customers that have a 'clunker' that is worth less than US$4,500 can benefit from participating, and therefore the number of purchases resulting from the program is related to the number of these clunkers. Other determinants included in the econometric model are earnings, household default rate, number of credit cards, house prices, and demographic measures. The estimated equation is as follows[7] (p. 1125):

$$\frac{Autopurchases_{cm}}{Autopurchases\,c_0} = \alpha^m + \beta^m \star CARS_c + \Gamma^{\mathrm{m}} \star \mathrm{Controls_c} + \varepsilon_{\mathrm{cm}}$$

Where:

$Autopurchases_{cm}$ = the number of cars sold in month *m* in city *c*

$Autopurchases\,c_0$ = the number of cars sold in the base month in city *c*

α^m = the effect of the month

β^m = the partial correlation between the number of clunkers in *c* and the endogenous variable

$CARS_c$ = the number of clunkers in *c*

Γ^{m} = the partial correlation between other explanatory variables and the number of purchases

The study (until this point) could equally plausibly be interpreted as presupposing the regularity view (if the econometric model was used for establishing a regularity between some variables of interest and car sales), or even a version of the manipulationist stance (considering that the estimation may be shown to rely on the design of a natural experiment). However, Mian and Sufi (2012) used the estimates obtained by running the specified equations to calculate the number of cars that would not be purchased if the 'cash for clunkers' program had not been introduced. To do so, they summarized the number of cars purchased in cities with exposure to the program and deduced the estimated number of cars purchased in the 'control group' consisting of the bottom decile of cities. Based on these estimations, Milan and Sufi concluded that in the short term, "340,000 cars were purchased under the program that would have otherwise not been purchased" (p. 1130). However, the number of cars sold in the subsequent months fell significantly when the 'cash for clunkers' program ceased to operate, and therefore this effect of the stimulus package was 'borrowed' from near future car sales rather than constituting an additional purchase of durable goods. Overall, the authors concluded that they found "no evidence of an effect" (p. 1106) of the 'cash for clunkers' program on either car sales or the employment and housing markets.

Furthermore, Mian and Sufi's (2012) claim instantiates a Galilean counterfactual (according to the distinction between Galilean and manipulationist counterfactuals introduced previously). Let me argue as follows. Analyzing the cross-sectional differentiation of the effects of an intervention (the purpose of the study under consideration) and predicting the effects of a planned manipulation should be distinguished. The control group in Mian and Sufi's (2012) study is a group of cities with the proportionally lowest number of clunkers. Therefore, their study is aimed at estimating the level of sales under the counterfactual scenario that other consumers from other cities are not eligible to take part in the stimulus program and is spatiotemporally limited to estimating the results of the specified stimulus. In other words, the study addresses the counterfactual question of what would be the number of cars sold in the United States if everything else were constant, but the program had not been introduced. Such a purpose differs from a study aiming to assess the effects of a specified intervention in being specific to a particular socioeconomic environment at a given time. The results cannot be extrapolated beyond a particular program that was at work in one country as a response to the 2007–2008 recession. Furthermore, the study is based on the potential outcome methodology that is interpreted as employing the counterfactual notion of causality.

The last step of the study indicates that Mian and Sufi (2012) presuppose the counterfactual view on causality. Rather than producing their causal conclusion on the grounds of the regularities found by means of econometric modeling, they calculated the results of the counterfactual scenario and estimated what car sales would be in each city if the number of registered clunkers equaled the figure of the control group (i.e., 10% of cities with the lowest number of clunkers).

The relata of the causal claim put forward by Mian and Sufi (2012) are variables. These variables (number of clunkers, number of purchased cars, housing prices, etc.) are features of aggregate-level phenomena instead of a quantitative description of events. Comparing to other econometric studies discussed previously, another distinct feature of this analysis is the focus on singular, token-level causal relations rather than general, type-level one. The estimation of the effects of cash for clunkers program is a representative example of drawing singular causal claims from a previously established econometric model.

4.2.2 Case study of the influence of intellectual property on R&D

As I have admitted previously, mainstream economists rarely employ case studies to infer singular (token-level) causal claims. However, this observation does not equal to saying that case studies are absent from all economic sciences. The difference results from the fact that the less mathematized economic sciences (management being the prime example) make use of this approach to causal inference. One of the few case studies that can be found among top economics papers is Heidi Williams' (2013) analysis of the influence of intellectual property rights (IP, henceforth) on research and development. The case study is aimed at estimating the effect of IP owned by Celera, a biotechnology company that partially sequenced the human genome, on the intensity of research and the number of products delivered to the market by pharmaceutical companies. In other words, Williams' purpose was to produce a token-level causal claim regarding the effects of IP based on an in-depth study of a single case. To obtain this goal, Williams (2013) has estimated three simple linear econometric models[8] and interpreted the parameters under interest counterfactually; i.e., establishing the Galilean counterfactuals.

The gene sequencing is a random process in the sense that the choice of a gene for sequencing is not a decision of a researcher. On the contrary, DNA is randomly cracked into smaller parts, coding single proteins (e.g., Tropp 2008). Furthermore, two research projects had aimed at sequencing the human genome. On the one hand, the Human Genome Project, a publicly funded consortium of research teams, put efforts into sequencing between 1990 and 2001, when the project was announced to be completed. On the other hand, Celera sequenced a selection of genes between 1999 and 2001 and protected them with a contract-law based IP (McBride 2002). Williams' (2013) study employs the assumption that the difference between the marketability and scientific potential of the genes sequenced by the two entities can be ascribed to the IP used by Celera.[9] To put it differently, the paper establishes that intellectual property is a token-level necessary condition for lower market and scientific use of genes.

Even though the paper's main conclusion that "Celera's IP led to reductions in subsequent scientific research and product development on the order of 20–30 percent" (p. 24) is not a counterfactual claim, counterfactual semantics are used intensively throughout the article. For example, Williams (2013, p. 16) concluded on the basis of a statistical comparison of the scientific outputs devoted to studying genes under Celera's IP law and being in the public domain that "if Celera genes had counterfactually had the same rate of subsequent innovation as non-Celera genes, there would have been 1,400 additional publications between 2001 and 2009 and 40 additional tests as of 2009." In other words, the case study establishes that Celera's IP is a necessary condition for reduced research output devoted to studying the genes under the IP law. Given the research design that focuses on a single case, the counterfactual is a token-level causal claim that partially describes the causal structure that produced the case study under analysis. Therefore, despite accepting that causes are necessary conditions, contrary to cliometric studies (cf. Section 2.2), the purpose of Williams' (2013) analysis was to establish a causal claim that is true only within the case under analysis and has severe implications for employing results of this type for policymaking. The relata of the causal claim are events (e.g., Celera using IP law to profit from sequencing their genes).

I must admit that this case is not only an exceptional use of the case study methodology in mainstream economics, but it also diverges from the usual approach to a case study analysis in its heavy reliance on econometrics. Possibly, the intensive use of econometric modeling allowed this nonstandard research technique to enter one of the leading economics journals. Another plausible reason for the publication of the paper in one of the top economics journals is the lack of systematic (in comparison to case-based studies) research on the adverse effects of intellectual property laws on innovativeness and research. The standard neoclassical economic theory predicts that IP should not hinder technological development, given that the assumption of no transaction costs is fulfilled (e.g., Green and Scotchmer 1995). The evidence opposing this mainstream view is scarce, and probably, considering the nature of the problem, more extensive cross-sectional studies are unlikely to appear due to the lack of cross-sectional data.

While the main result that there is a systematic difference (that can be ascribed to the IP) in use of genes sequenced by Celera (and being its IP) and the publicly funded Human Genome Project is established by simple statistical comparisons of averages between the two samples presented in Table 1 (see Williams 2013, p. 4), three different econometric models corroborate this result with a view to establish that it is the IP law that impedes research and commercial use of these genes instead of other factors such as, for example, Celera's genes being systematically less marketable and even less noteworthy for the biotechnology community (getting fewer citations). However, the conclusion of the study is limited only to establishing a token-level counterfactual describing the necessary condition given other causal factors that operated within the case study.

Williams (2013, p. 14 et seq.) controlled for other influences such as the number of publications in a given year, year of disclosure, and molecular covariates in a simple OLS linear regression. Williams' second empirical test can also be interpreted as a regression discontinuity design (more on that topic in Chapter 6). If the emphasis were put on the influence of waiving IP on the citability of each gene, then the panel estimates of the number of citations before and after Celera's

IP has been waived (because the Human Genome Project sequenced a given gene). However, considering that this regression is a part of a more extensive study aimed at establishing the counterfactual token-level causal claim depicting the influence of Celera's IP on research and innovation, then this regression can be interpreted as a robustness check instead. Finally, the third regression studies the correlation between the time of a gene being under IP and its citability and use in DNA tests. The number of innovation outcomes is linearly related to the time a gene is in public domain, and the earlier a gene's use stopped being hampered by IP law, the more innovative products using that gene have been marketed (cf. 21). Given these results, Williams' (2013) study establishes that if the IP law were not operating, the genes sequenced by Celera would be used more often in both research and market products.

4.3 Counterfactuals and economic policymaking

Contrary to other views on causality, the counterfactual approach overlaps with other philosophical views on what the relation is between causes and effects. It employs the 'causes as necessary conditions' view, which is one of the definitions coined on the grounds of the regularity view. However, counterfactuals can also be entangled with manipulationist theories. For example, Woodward (2005) formulated his manipulationist theory of causality employing counterfactual claims (cf. Chapter 6). By analogy, discussing the effects of a stimulus package is obviously connected to the manipulationist view on causality. Evidence supporting a counterfactual claim describing the effects of Woodward's intervention needs to be broader than evidence for a claim describing the presence of a necessary condition in a particular phenomenon under study. Unfortunately for the use of causal claims established by the studies considered in the previous section, the economists put forward Galilean counterfactual token-level causal claims, and therefore its use for policy requires extrapolating the result from the case under consideration into a policy setting. Considering that the problem of extrapolation is discussed in more depth in Section 6.3, here, I briefly analyze what additional knowledge is required for extrapolation to be fruitful for policy and conclude pessimistically by repeating Cartwright and Reiss' (2008) conclusion that policymakers should only use policy-relevant counterfactuals, but such counterfactuals can virtually never be found in published research. I consider the challenges connected to gathering evidence for such claims based on the case of discussion of the Troubled Assets Relief Program that was introduced in the Autumn of 2008 in the United States.

Both Mian and Sufi's (2012) and Williams' (2013) causal claims are Galilean counterfactuals that address the question of what would happen if *C* (i.e., the 'cash for clunkers' program or Celera's IP) did not happen. Given that these two studies establish token-level causal claims that only describe the phenomena under study, they do not have direct policy implications. For example, if a policymaker wanted to use Mian and Sufi's (2012) evidence to introduce similar (or even the same) program in Poland, then they would have to establish (on the basis of other studies) that the causal structure underlying the car market in the United States and Poland is the same. In other words, justified extrapolation requires knowledge of other necessary and sufficient conditions that make such a program efficient. Without this additional knowledge, (unjustified) extrapolation can lead to unexpected outcomes. For example, considering that the cars on the Polish streets are, on average, older compared to their American counterparts, such a program could be more popular, especially considering the informational asymmetry between car owners and buyers, and the under average price level on the Polish used car market compared to other European countries (Akerlof 1978).[10]

Similarly, Williams (2013) delivers evidence that the presence of IP reduced both scientific use and marketability of technology (sequenced genes) owned by Celera. However, her study (taken on its own) does not deliver evidence for generalizing this claim to other technologies. It is possible that other technologies (e.g., being more profound for marketable products or traded on markets with diminishable transaction costs), if they were subject to IP law, were not impeded by this fact. Unfortunately for the policy implications of case studies, these considerations show that establishing that a condition is a necessary cause for an effect does not equal to having evidence for it to be a sufficient cause. In other words, the evidence that C caused E given background conditions B_1 does not warrant that implementing C will result in E regardless of different background conditions B_2. Therefore, neither the Galilean counterfactuals established on the basis of econometric models nor the causal claims resulting from case studies suggest that policy resulting from their direct applications will be successful. To employ the causal claims based on case studies, a policymaker needs to establish that all other conditions producing the phenomenon under study are present within the policy context. Unfortunately, this task requires additional knowledge that would likely suffice for implementing a policy on its own.

Cartwright and Reiss (2008) argued that policymakers should search for the counterfactual causal claims that are relevant for the interventions under consideration. While their concluding remarks are of general nature, their advice can be interpreted as using as evidence for policymaking only counterfactuals of the following form: given background conditions B_C, intervention I results in effect E. However, these types of 'policymaking-oriented counterfactuals' are scarcely available in the existent literature, and therefore this conclusion advises policymakers to gather evidence and establish context-specific counterfactuals on their own.

However, putting forward the manipulationist counterfactuals is an epistemically demanding task. Let me consider the introduction of the Troubled Asset Relief Program (TARP) on 3 October 2007, at the outbreak of the 2007–2008 financial crises (Herszenhorn 2008). At the time, the housing market bubble collapsed because the spike of interest rates made homeowners struggle to meet their mortgage payments. Consequently, mortgage-backed securities (MBSs) such as collateralized debt obligations (CDOs), despite obtaining high credit ratings, proved to be worth only a small fraction of their book value. The crisis was additionally worsened by high leverage used by financial institutions to raise potential profits from investments that were considered very safe. After a few internationally recognized companies became insolvent due to the excessive risk-taking on the MBS market and swaps on interest rates, in the second half of 2008, the interbank short-term loan market ceased to exist. Given that detailed lists of assets owned by companies are not publicly available information and therefore only top management of financial institutions was aware of their exposure to the toxic MBS, market participants were unable to assess whether investment banks and other financial institutions willing to borrow money on a short term are going to stay solvent in the following days and weeks (Brunnermeier 2009). The frozen short-term inter-bank credit market was likely to negatively affect the accessibility of financing for other companies, and, consequently, a significant recession resulted from insufficient demand.

The bailout was mainly spent on purchasing privileged stocks of financial institutions and buying MBSs from the institutions at the risk of insolvency. Its introduction was preceded by harsh debates in the House of Representatives and Senate. These debates allow us to have a glimpse into what the counterfactual causal claims used by policymakers look like, but they also indicate how difficult it is to foresee the effects of interventions. The effects of TARP on the economy have been discussed by Treasury Secretary Henry Paulson in the US Senate (Treasury 2008). His speech employs causal language from the very beginning, as Paulson admitted that previous stimulus actions such as the nationalization of Fannie Mae and Freddie Mac "have been necessary but not sufficient."

Further, Paulson argued that TARP targets the leading cause of the market turmoil, and therefore its introduction is a sufficient cause for prosperity. Therefore, his talk can be summarized with the following manipulationist counterfactual: 'if TARP were introduced, the deep economic recession would not happen.' While the detailed evidence leading to developing TARP stays unknown, we may suspect that the economists working for the US Treasury analyzed counterfactual scenarios on econometric models and/or ran simulations of how market players behave when toxic financial instruments disappear from the market. In particular, TARP was aimed at improving the liquidity of the short-term credit market and lowering the risk aversion of financial institutions. These steps should result in raised (from the recessionary level) global demand and, accepting the standard Keynesian model, improved economic development.

In a similar manner, Ben Bernanke, then the chairman of the Federal Reserve, delivered a similar rationale for the bailout and indicated that the implementation of TARP would result in (1) reducing risk aversion on financial markets, and (2) promoting GDP growth (Federal Reserve 2008). The counterfactual causal claims that convinced the US Senate to vote for the bailout proved true. According to Berger et al. (2016) study using a difference-in-differences approach (discussed in Chapter 6), TARP reduced systematic risk of the financial sector. Furthermore, the bailout increased job creation and other measures of economic prosperity (Berger and Roman 2017). However, the effects of the bailout were not only positive.

On the contrary, TARP resulted in what could be called side effects if the program were a drug. One of the most severe adverse effects on the economy were bank runs that occurred after the program was introduced. A possible explanation for the timing of the two events is the signaling role TARP played: bailout plans show that governments are aware of the bad debt owned by financial institutions (Wang 2013).

Overall, the positive effects of TARP (mainly preventing a more profound economic depression) seem to outweigh its detrimental results, and therefore the bailout can be used as a case of a successful intervention. Even though interventions can in principle be successful despite inappropriate evidence in their favor, it is probably more likely that successful policy actions are based on plausible evidence. What follows, Ben Bernanke and Henry Paulson's manipulationist counterfactuals describing the effects of introducing TARP, are likely based on evidence rightly identifying other necessary conditions for improving economic development in the aftermath of the 2007–2008 financial crisis. However, the views of economists were divided. As Herszenhorn (2008) put it, "[s]upporters said the bailout was needed to prevent economic collapse; opponents said it was hasty, ill-conceived and risked too much taxpayers' money to help Wall Street executives while providing no guarantees of success." The division of opinions on the manipulationist counterfactual regarding the effects of bailout show how demanding it is assessing the claims describing the effects of considered interventions. Putting forward well-justified manipulationist counterfactuals requires running counterfactual analyses of models and presupposes intervention-invariance of represented relationships. The process involves several decisions regarding the choice of a relevant econometric model, its estimation, and finally running a counterfactual scenario. All these decisions are, to some extent, subjective in the sense that different choices are plausible to explain vast disagreements in economists' opinions regarding effectiveness of policies. Similarly to statistical techniques used to aggregate medical evidence (Stegenga 2018, p. 78), the methods of causal inference used in economics are malleable. Otherwise, the differences in opinions regarding economic policymaking, fueled sometimes by political linings, could be resolved by gathering more evidence and correcting methodological flaws.

Without a reliable and relevant manipulationist counterfactuals, extrapolating from a case study into a target (policy setting) requires obtaining knowledge (more on this topic in the following chapters) of the target. That knowledge, by itself, would allow for policymaking, because case studies do not deliver evidence for deciding whether the background conditions present in the phenomenon under the case study operate within the policy target. However, the Galilean counterfactuals, despite being insufficient evidence for policymaking, can serve the purpose of assessing previous policy interventions.

4.4 Counterfactuals for the sake of knowledge or policymaking?

For Hume (1992, p. 56), the counterfactual causality serves as an ontological definition of causality, while the constant regularities are a hint of their presence. To the contrary, economists mostly employ this notion of causality as a secondary concept, to put forward singular causal claims. Mian and Sufi's (2012) assessment of the effectiveness of the cash for clunkers program is a prime example of an analysis establishing a counterfactual causal claim based on a previously established model. To do so, economists first estimate an econometric model representing a plausible causal structure of a phenomenon under study (both the causal relation under analysis and other background conditions), and second consider what would happen if one of the causal factors were different. Williams' (2013) case study of the influence of intellectual property on the scientific and practical use of technology instantiates another approach to establishing counterfactual causal claims. Even though this method of causal inference is exceptional for mainstream economics, it is often used in the less mathematized social sciences such as management. These two types of evidence for counterfactual causal claims define causes as necessary conditions and address the question of what would have happened if an event (cause) under consideration had not occurred.

The causal claims of this type are not sufficient evidence for policymaking. Extrapolating the counterfactual defining causes as necessary conditions from a case study into a policy setting requires policymakers to establish that all other necessary conditions for an effect to occur are present within the policy setting. Unfortunately, neither counterfactual analysis of econometric models nor case studies deliver evidence regarding the causal factors that operate *within* a policy setting. Therefore, while the research aiming at establishing the Galilean counterfactual causal claims are fruitful in assessing the effectiveness of policy interventions *ex-post*, they cannot be directly applied to intervening. To do so, a policymaker needs to establish that other necessary causes (background conditions) producing a phenomenon under study are present within the policy context. Fulfilling this task would establish a manipulationist counterfactual describing the effects of a specific intervention. Unfortunately, it would require possessing broad knowledge of the causal relations present in the policy context.

Notes

1 Here, I differentiate between academic economic research and the work as experts, and refer to the former.
2 Galilean counterfactuals bear their name from Galileo Galilei because of their overlap with the Galilean idealization. Similarly to the law of falling bodies, Galilean counterfactuals describe relations idealized by omission. Therefore, while they can serve the purpose of addressing 'what if' questions under the *ceteris paribus* clause, their direct application to predicting the outcomes of intervention may lead to fallacious results.
3 Heidi's study is exceptional for contemporary mainstream economics (Maziarz 2018). However, discussing this study may be beneficial considering the popularity of the qualitative methods in related disciplines such as management.
4 The value of payments depended on the fuel efficiency of cars traded in by consumers in exchange for a reduced price of newer cars, and ranged from US$3,500–4,500.
5 Mian and Sufi (2012, p. 1115) used dataset covering 957 US metropolitan and micropolitan statistical areas (core-based statistical areas, or CBSAs). Therefore, the notion of 'city' refers not only to actual cities, but also to rural areas.

6 R.L. Polk categorizes vehicles into groups covering a few models of cars for each year. Due to the imperfect math between the two datasets, proceeding further required some degree of approximation.

7 In addition to the basic specifications listed, Milan and Sufi estimated several other specifications (e.g., using a difference in logarithms of the number of cars purchased, estimating a cumulative version of the equation or excluding other controls) with a view to check the robustness of the finding.

8 Even though case studies, by definition, focus on analyzing single instances, estimating econometric models is possible in this case because Celera owned patents on sequenced genes that are *ex ante* homogenous in regard to their marketability or being fruitful in research. For this reason, Williams' (2013) study is extraordinary in its reliance on econometric modeling.

9 The random allocation of genes between the IP-protected and public-domain entities could also have been interpreted as a quasi-experimental research design (see Chapter 6). However, the counterfactual analysis and semantics that appear throughout the paper suggest a different reading.

10 A lower price of lemons or cars (the goods whose quality is known to the owner, but not to the buyer) in particular markets indicates a lower average quality of the traded goods. If owners know the quality of the car is lower than average, they have financial incentives to take part in a program similar to the 'cash for clunkers' stimulus package.

References

Akerlof, G. A. (1978). The market for 'lemons': Quality uncertainty and the market mechanism. In: *Uncertainty in Economics* (pp. 235–251). Academic Press. DOI: 10.1016/B978-0-12-214850-7.50022-X

Becker, G. (2010). The cash for clunkers program: A bad idea at the wrong time. *Becker-Posner Blog*. Retrieved from: www.becker-posner-blog.com/2009/08/the-cash-for-clunkers-program-a-bad-idea-at-the-wrong-time-becker.html. Access: 18th February 2019.

Bennett, A. (2004). Case study methods: Design, use, and comparative advantages. In: *Models, Numbers, and Cases: Methods for Studying International Relations* (pp. 19–55). Ann Arbor: University of Michigan Press.

Berger, A. N., & Roman, R. A. (2017). Did saving wall street really save main street? The real effects of TARP on local economic conditions. *Journal of Financial and Quantitative Analysis*, *52*(5), 1827–1867. DOI: 10.1017/S002210901700062X

Berger, A. N., Roman, R. A., & Sedunov, J. (2016). Do bank bailouts reduce or increase systematic risk? The effects of TARP on financial system stability. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2708953. Access: 21st June 2019.

Blinder, A. (2008). A modest proposal: Eco-friendly stimulus: Economic view. *New York Times*, 27th July.

Broadbent, A. (2013). Why philosophy of epidemiology? In: *Philosophy of Epidemiology* (pp. 1–9). London: Palgrave Macmillan.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives*, *23*(1), 77–100. DOI: 10.1257/jep.23.1.77

Bunkley, N. (2010). Incentives foster a major jump in auto sales. *New York Times*, 1st April.

Cartwright, N. (2007). Counterfactuals in economics: A commentary. In: *Causation and Explanation* (pp. 191–216). Cambridge: MIT Press.

Cartwright, N., & Reiss, J. (2008). Uncertainty in econometrics: Evaluating policy counterfactuals. In: Reiss, J. (ed.) *Error in Economics: Towards a More Evidence-Based Methodology* (pp. 187–213). London: Routledge.

Dawid, A. P. (2000). Causal inference without counterfactuals. *Journal of the American Statistical Association*, *95*(450), 407–424.

Dion, D. (2003). Evidence and inference in the comparative case study. In: *Necessary Conditions: Theory, Methodology, and Applications* (pp. 95–112). New York: Rowman and Littlefield Publishers Inc.

Dul, J., & Hak, T. (2008). *Methodology in Business Research*. Amsterdam, Boston, Heidelberg: Butterworth-Heinemann.

Fearon, J. (1991). Counterfactuals and hypothesis testing in political science. *World Politics*, *43*(2), 169–195. DOI: doi.org/10.2307/2010470

Federal Reserve (2008). Testimony: US financial markets. Retrieved from: www.federalreserve.gov/newsevents/testimony/bernanke20080923a1.htm. Access: 19th June 2019.

Frisch, M. (2005). *Inconsistency, Asymmetry, and Non-Locality: A Philosophical Investigation of Classical Electrodynamics*. New York: Oxford University Press.

Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, *98*(2), 341–354. DOI: 10.1017/S0003055404001182

Gerring, J. (2006). *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Goertz, G., & Levy, J. S. (2007). Causal explanation, necessary conditions, and case studies. *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals* (pp. 9–45). London: Routledge.

Green, J., & Scotchmer, S. (1995). On the division of profit in sequential innovation. *RAND Journal of Economics*, *26*(1), 20–33.

Hausman, D. (1998). *Causal Asymmetries*. Cambridge: Cambridge University Press.

Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century retrospective. *The Quarterly Journal of Economics*, *115*(1), 45–97. DOI: doi.org/10.1162/003355300554674

Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, *76*(1), 1–27. DOI: 10.1111/j.1751-5823.2007.00024.x

Henschen, T. (2018). What is macroeconomic causality? *Journal of Economic Methodology*, *25*(1), 1–20. DOI: 10.1080/1350178X.2017.1407435

Herszenhorn, D. M. (2008). Bailout plan wins approval; Democrats vow tighter rules. *New York Times*. Retrieved from: https://www.nytimes.com/2008/10/04/business/economy/04bailout.html. Access: 18th October 2018.

Hume, D. (1992). *An Enquiry Concerning Human Understanding*. London: Hackett Publishing.

King, G., & Zeng, L. (2007). When can history be our guide? The pitfalls of counterfactual inference. *International Studies Quarterly*, *51*(1), 183–210. DOI: 10.1111/j.1468-2478.2007.00445.x

Lewis, D. (1973). Causation. *The Journal of Philosophy*, *70*(17), 556–567. DOI: 10.2307/2025310

Lewis, D. (1979). Counterfactual dependence and time's arrow. Noûs, *13*(4), 455–476. DOI: 10.2307/2215339

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, *61*(4), 343–377. DOI: 10.1080/00048408312341131

Lewis, D. (2000). Causation as influence. *The Journal of Philosophy*, *97*(4), 182–197. DOI: 10.2307/2678389

Lewis, D. (2004). Void and object. In: Collins, J. D., Hall, E. J., & Paul, L. A. (eds.) *Causation and Counterfactuals*. New York: MIT Press.

Little, R. J., & Rubin, D. B. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. *Annual Review of Public Health*, *21*(1), 121–145. DOI: 10.1146/annurev.publhealth.21.1.121

Mackie, J. (1974). *Cement of the Universe*. Oxford: Clarendon Press.

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly*, *2*(4), 245–264.

McBride, M. S. (2002). Bioinformatics and intellectual property protection. *Berkeley Tech. LJ*, *17*, 1331.

Maziarz, M. (2018). Causal inferences in the contemporary economics. *Mendeley Data*. Retrieved from: http://dx.doi.org/10.17632/v7dhjnd8xg.2. Access: 16th October 2018.

McCloskey, D. N. (1998). *The Rhetoric of Economics*. Madison: University of Wisconsin Press.

Menzies, P. (2017). An interventionist perspective. In: Beebee, H. et al. (eds.) *Making a Difference: Essays on the Philosophy of Causation* (pp. 153–174). Oxford: Oxford University Press.

Mian, A., & Sufi, A. (2012). The effects of fiscal stimulus: Evidence from the 2009 cash for clunkers program. *The Quarterly Journal of Economics*, *127*(3), 1107–1142. DOI: 10.1093/qje/qjs024

Morgan, S., & Winship, Ch. (2014). *Counterfactuals and Causal Inference*. Cambridge: Cambridge University Press.

Morini, M. (2011). *Understanding and Managing Model Risk: A Practical Guide for Quants, Traders and Validators*. London: John Wiley and Sons. DOI: 10.1002/9781118467312

Paul, L. A. (2009). Counterfactual theories. In: Beebee, H. et al. (eds.) *The Oxford Handbook of Causation*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199279739.003.0009

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Oxford: Oxford University Press.

Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press.

Reiss, J. (2015). *Causation, Evidence, and Inference*. London: Routledge. DOI: 10.4324/9781315771601

Rubin, D. B. (1990). Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, *5*(4), 472–480.

Salmon, W. (2006). *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.

Shpitser, I., & Pearl, J. (2012). Effects of treatment on the treated: Identification and generalization. *arXiv preprint arXiv:1205.2615*.

Stegenga, J. (2018). *Medical Nihilism*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198747048.001.0001

Strawson, G. (1987). Realism and causation. *The Philosophical Quarterly (1950–)*, *37*(148), 253–277. DOI: 10.2307/2220397

Treasury (2008). Testimony by Secretary Henry M. Paulson, Jr. before the Senate Banking Committee on Turmoil in US Credit Markets: Recent Actions regarding Government Sponsored Entities, Investment Banks and other Financial Institutions. Retrieved from: www.treasury.gov/press-center/press-releases/Pages/hp1153.aspx. Access: 18th June 2019.

Tropp, E. (2008). *Molecular Biology: Genes to Proteins*. Boston: Jones and Barlett Publishers.

Wang, C. (2013). Bailouts and bank runs: Theory and evidence from TARP. *European Economic Review*, *64*, 169–180. DOI: 10.1016/j.euroecorev.2013.08.005

Williams, H. L. (2013). Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, *121*(1), 1–27. DOI: 10.1086/669706

Woodward, J. (2005). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

5 Mechanisms

Usually, causal-process theories are listed as one of the big five approaches to causality. Phil Dowe (2012) defined this family of theories in the following way: "[a]ccording to the process theory, any facts about causation as a relation between events obtain only on account of more basic facts about causal processes and interactions. Causal processes are the world-lines of objects, exhibiting some characteristic essential for causation" (p. 213). This approach was supported by Wesley Salmon (1984), who put forth arguments according to which only causal explanations are explanatory. According to his stance, knowing the correlation between money and prices cannot be used for the purpose of explanation unless one assures that the correlational dependence is produced by a relation of causal nature; or, to put it differently, by a causal mechanism connecting the two. Wesley Salmon (1984) indicated that processes have constant structure over time and transmit modifications. The conserved-quantity theory, which is a traditionally held viewpoint among physicists, states that actual processes transmit energy (Krajewski 1997) and it is the feature that distinguishes causal and noncausal relations. These and similar theories all have in common the emphasis on physics and natural sciences. Therefore, they may be incongruent with the social sciences.

On the one hand, causal-process theories can be interpreted as a special case of the mechanistic approach to causality (Glennan 2012). On the other hand, Jon Williamson (2011a) believed that the causal-process and mechanistic approaches are equal subgroups of a broader realistic approach to causality. He distinguished between process theories that suit the causal analysis focused on physics and complex-systems theories. The former group of theories states that "*A* causes *B* if and only if there is a physical process of the appropriate sort that links *A* and *B*" (Williamson 2011a, p. 423). The latter family of theories defines causal relations in terms of being parts of the same complex-systems mechanism, where "[a] complex-systems mechanism is a complex arrangement of parts that is responsible for some phenomenon partly in virtue of the organization of these parts." Taking into account the relevance of the two approaches to economic research, the chapter focuses on the family of mechanistic theories of causality. Furthermore, 'causal process' and 'mechanism' are sometimes considered synonymical on the grounds of the philosophy of social science (Maxwell 2004; Collier 2011).

The mechanistic approach to causality identifies causal relations with mechanisms underlying or producing effects. The notion of mechanism is employed "to refer to a variety of systems or processes that produce phenomena in virtue of the arrangement and interaction of a number of parts" (Glennan 2012, p. 315). Ontologically, causal relations are identified with mechanisms constituted by causes that produce effects. Epistemically, accepting this approach leads to accepting only such inferences about causal relations that deliver mechanistic evidence. The two notions seem to be interrelated in the research practice of economists.

The first part of Section 5.1 focuses on summarizing the views of philosophers of causality on what mechanisms are and what role do they play in causal inference. I also review the discussion of causal mechanisms in the philosophy of economics. In Section 5.2, I exemplify using theoretical modeling as a method of conjecturing about economic mechanisms with a discussion of Matthew Kotchen's (2006) purely theoretical model of markets for public goods, Ian Parry and Kenneth Small's (2009) calibrated theoretical model of the influence of subsidies on public transport use, and Lawrence Christiano et al.'s (2011) DSGE model. In Section 5.3, I focus on the use of mechanistic knowledge for policymaking and argue that economic models often represent possible mechanisms and it is policymaker's duty to establish the actuality of a mechanism under consideration. Unfortunately for policymakers, the evidence needed to establish the actuality of a mechanism is sufficient for producing a causal claim on its own, what I label as 'mechanist's circle' due to the similarity of the problem with extrapolating the effects of experiments considered in the following Chapter 6 known as extrapolator's circle. Furthermore, the knowledge of one mechanism is not sufficient for predicting the effects of interventions modifying the relata of causal claims due to the influence of other mechanisms that are likely to make the outcomes differ from the expectations based on the study of one mechanism. However, mechanistic evidence can be employed for policymaking relying on designing mechanisms; i.e., institutional reforms.

5.1 The mechanistic theories of causality

The mechanistic approach to causality seems to become a hot topic in philosophical debates. Bechtel and Richardson (2010, p. xvii et seq.) foretell the appearance of a new, mechanistic paradigm in the philosophy of science. Apart from the debates on causality, philosophers employ the notion of mechanisms to solve the problem of extrapolation in natural and social sciences, and argue for raising the importance of mechanistic evidence in the projects of evidence-based medicine and policy. According to Stuart Glennan (2002), mechanisms were introduced to the contemporary debate by Peter Railton (1978), who coined a model of statistical explanation competing with Carl Hempel's (1965) inductive-statistical model. Facing the problem of rare (nomically unexpectable) events that also seek an explanation despite being in disagreement with Hempel's requirement that explanandum must be probable, Railton demanded from explanations to contain an account of mechanisms defined as "more or less complete filling-in of the links in the causal chains" (p. 748). Glennan (2002) distinguished between the Salmon–Railton approach and mechanisms-as-complex-system definition. The former way of defining mechanisms puts emphasis on the step-by-step process (chain) of events. An exemplification of the former type of mechanism is the following process: (1) nudging a glass of coffee, (2) its falling from a table, and (3) glass shattering on a concrete floor. The latter mechanisms depict the operation of complex systems such as a watch or social interactions within a company. Phyllis Illari and Jon Williamson (2012) indicated that a fierce debate on the meaning of mechanisms occurred during recent decades. Today, the realistically minded philosophers of causality have not yet established a consensus. The term is ambiguous and variously defined across different fields and by various theorists. Following, a few notable definitions are reviewed.

5.1.1 What is a causal mechanism?

Peter Machamer et al.'s (2000) definition of mechanism entered the philosophy of economics by Caterina Marchionni's (2017) work addressing the question of what is an economic mechanism. Their definition states that mechanisms are "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions." (Machamer et al. 2000, p. 3). Moreover, they highlighted that one of the features of mechanisms is that they are regular; i.e., they produce the same phenomena in the same conditions. In other words, the presence of a mechanism guarantees the reproducibility of experimental results. However, as I argue ahead, the reproducibility holds only within a particular experimental setting. In an open world, to employ Lawson's (1997) label, where several counteracting mechanisms operate (some of which are likely to stay unknown), the effects of its activity are not certain unless one knows the rules describing how they interact. Another seminal definition reduces mechanism to "structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena" (Bechtel and Abrahamsen 2005, p. 423). In general, the term 'mechanism' denotes either entities interacting with each other or a process producing effect. Machamer et al. (2000, p. 5) differentiate between substantivist and process ontologies presupposed in the definitions of mechanisms. The former approach focuses on analyzing entities and their properties. The latter identifies mechanisms with processes and limits the role of entities. The dualist consensus according to which these two approaches are in agreement and depict different settings is establishing (Glennan 2012).

According to James Wright (2012, pp. 376–377), the notion of mechanism is essential for a group of contemporary philosophers of science called 'New Mechanists' who define the process of explanation as the activity focused on exhibiting mechanisms that are responsible for (produce) causal patterns and regularities. New mechanists accept the definition of mechanism that indicates that (1) mechanisms are hierarchical systems, (2) parts of these systems are coordinated, (3) the coordination is responsible for producing a higher-level activity, and (4) this higher-level activity is identified as a phenomenon seeking an explanation. Bechtel and Richardson (2010, p. xviii) defined mechanistic explanation as the identification of "parts and their organization, showing how the behavior of the machine is a consequence of the parts and their organization."

Stuart Glennan (1996, p. 49) defined causal mechanisms as "complex systems whose 'internal' parts interact to produce a system's 'external' behavior." He later added (p. 52) that a causal mechanism is a "complex system which produces [. . . a] behaviour by the interaction of a number of parts according to direct causal laws." Finally, Glennan (1996, p. 56) delivered the definition of causal relations according to which *A* causes *B* when an appropriate mechanism connects the two events. Such a definition can be problematic in the context of economics. For example, the attempt to translate the classical Keynesian IS-LM model into a mechanism cannot stop at the level of microeconomics due to the lack of 'direct causal laws.' Macroeconomic phenomena are explained by microeconomic mechanisms; these should be explained by psychological theories, but even they are not nonreductive. Similar considerations lead Jon Williamson (2011a) to accuse the mechanistic approach of regress ad infinitum. At least in the case of social sciences, there are virtually infinitely many steps of regress. The problem of regress ad infinitum may not appear in the research practice. Lindley Darden[1] (2002, p. 356) defended mechanistic view on causality from such criticism indicating that "[f]or a given scientific field, there are typically entities and activities that are accepted as relatively fundamental or taken to be unproblematic for the purposes of a given scientist, research group, or field." In the case of economic mechanisms, such basic activities would probably include maximizing behavior and having preferences (Marchionni 2017). Contrary to Darden, James Woodward (2002), who is known for his manipulationist theory of causality (cf. Chapter 6), attempted at defining causal mechanisms using counterfactual conditions.

Craver and Tabery (2017) delivered a useful summary of the features of mechanisms and listed a few misconceptions of it. First, mechanisms can be both deterministic and indeterministic. Its indeterminism can be both an effect of being produced by nonstrict, statistical laws, and interference of outside factors. Second, mechanisms can result from the operation of both machines and complex systems that are not designed to produce it. Third, the causal laws produced by mechanisms can be linear or nonlinear, and have a sequential or more complex nature. Fourth, mechanisms can be either localized in a particular place (e.g., within a watchtower) or the complex system can be distinguished considering the connections between its parts (e.g., the participants in an online market for a futures contract that lack physical connections). Fifth, the causal connections between parts of mechanisms are not limited to physical relations. They can be exemplified with all kinds of causal relations employed by both natural and social sciences. Finally, mechanisms are not only more detailed (lower-level) theories that strive for delivering an explanation of a mechanism. There is an ontological dimension of the mechanistic approach to causality: mechanistic explanations are committed to acknowledging the existence of the complex systems they depict. Despite Craver and Tabery's (2017) claim that the appeal to mechanism can be nonreductionistic, they seem to be so in the sense that mechanisms underlie causal relations. Therefore, these relations observed on a given level of reality are reduced to a lower-level mechanism.

Furthermore, Craver and Tabery (2017) listed the following four features standard for virtually all definitions of mechanisms. First, a mechanism produces a phenomenon. Second, a mechanism is constituted by interacting parts. Third, these parts are organized characteristically. Fourth, the parts constituting a mechanism are related.

Moreover, the characteristic feature of mechanisms is that they produce a phenomenon. In other words, if a relationship were a random correlation that is not produced by (lower-level) entities and relations among them (there is no mechanism), then there is no causal relation. Finally, the fundamental causal relations that join parts of a mechanism are not mechanisms themselves. The last feature of mechanisms defends the mechanistic approach from the accusation of regress ad infinitum. Williamson (2011a) argued against the accusations of circularity in a different way. According to his stance, the causal law at a higher level (produced by a mechanism) is something different than the causal connections of elements of a complex system. The higher-level causal laws are regularities produced by a mechanism. On the contrary, the connections between elements of a complex system are not mechanistic in nature, but are more basic.

Accepting the definitions of mechanisms indicating that they produce the higher-level causal relations poses the question of what the relation is between mechanisms and the cause-and-effect relationships. Glennan (1996) put forth the argument that what is observable as a causal relation on a particular level of reality is a mechanism on an underlying level. For instance, the law of demand can be described as a causal relation between price P and quantity Q of a good sold on a market. On a lower level of reality, the law of demand is produced by many consumers (entities) undertaking decisions regarding the quantity Q of a good they want to buy at a particular price level P that are themselves mechanisms seeking an explanation.

5.1.2 What are the mechanisms suitable for?

I have argued that data-driven causal inferences are susceptible to the common-cause fallacy. Correlational studies are sometimes accused of mistaking actual causal relations with spurious dependencies. The most significant advantage of identifying causality with mechanisms is getting a straightforward demarcation criterion among truly causal connections produced by mechanisms and accidental regularities (cf. Moneta and Russo 2014). Glennan (2012) indicated that understanding a causal mechanism that binds X and Y makes it possible to predict when (in what circumstances) such a relation will break down what proves very useful for policymaking. While I argue that the complication of the social world undermines such views concerning economics, they possibly apply to more advanced sciences or analyzing less complicated fields. Mechanistic evidence is usually considered as the real gold standard (contrary to experimentation that is accused of being unjustifiably labeled in this way (cf. Cartwright 2007). Such an opinion is, to some degree, justified. A seminal example in favor of this viewpoint is to consider the relationship between crime rate and police patrols. Wilson and Kelling (1982) indicate that the (observed) number of crimes in a particular area raises if number of police patrols is increased. Employing a probabilistic or manipulationist accounts of causality does not enlighten beyond observed phenomena, and, according to these approaches, police patrols cause crime. However, mechanistic evidence indicates that increasing number of police patrols positively influence observed quantities, while the 'real' number of crimes can be in fact reduced. Till Grüne-Yanoff (2016) argued that describing an actual mechanism makes plausible predicting whether a policy will prove useful. Despite its cheers, a significant issue with mechanistic models is the fact that their truth (i.e., the resemblance of the mechanism grasped by model to the processes taking place in the real world) seems to be difficult to justify.

There are two dimensions of the mechanistic view on causality. Ontologically, mechanisms are entities and processes that produce observable phenomena. The regularity and probability approaches discussed previously face the dilemma of putting forth a distinction between causal and random associations. The usual solutions are either based on the belief that randomly produced regularities will vanish in the limit (i.e., when the number of observations is sufficiently high) or assume that only regularities that instantiate laws of nature are causal (cf. Chapters 2–3). The mechanistic approach to causality develops its own solution: only those regularities that are produced by a mechanism are of causal nature (Glennan 1996). In other words, causal claims about a considered level of reality are justified if a mechanism producing such relation exists at a lower level. Epistemically, mechanisms are theoretical descriptions of phenomena under consideration (mechanistic theory). Reiss (2007, p. 166) highlighted that the ontological and epistemic understandings of mechanisms are often entangled. In a similar vein, Joseph Maxwell (2004) indicated that discussing causal mechanisms is inseparable from being a realist about causal processes. In a similar vein, Machamer et al. (2000, p. 1) highlighted that a "satisfactory explanation" is delivered by theory depicting an actual mechanism. In other words, only true mechanistic theories explain. Unfortunately, the question of how to assert that a theory delivers a correct explanation (i.e., represents how a mechanism operates in the world) waits for a satisfactory answer.

However, the mechanisms can be refused to exist outside of our theories if one interprets them in line with a version of the instrumentalist perspective on science. For instance, according to the epistemic theory (Williamson 2011a, 2011b; Russo and Williamson 2007), an investigator should strive for a context-dependent explanation. Petri Ylikoski (2012) indicated that considering that, usually, many mechanisms produce a phenomenon under analysis, the level of a mechanism under consideration depends on the target of explanation; i.e., they are perspectival, not objective. As Glennan (2002) put it, "mechanisms are not mechanisms *simpliciter*, but mechanisms *for* behaviors. A complex system may exhibit several different behaviors, and the decomposition of the system will depend upon which behavior is under consideration" (p. 344, emphasis in original). According to this view, there is a subjective element in the process of inference of mechanisms.

Glennan (2010) distinguished between stable (robust) and ephemeral mechanisms. The former kind of mechanism is usually studied by the natural sciences (e.g., a clock, a chemical reaction). The latter kind can be exemplified with a market of a particular good, or internal interactions between employers of a company. The ephemeral mechanisms are often observable in the realm studied by social sciences. However, even the mechanisms operating in the realm of social sciences should characterize a certain degree of stability. For instance, Woodward (2002) argued that one of the crucial features of mechanisms is that they are stable (invariant under interventions). Such a viewpoint is also acknowledged in the philosophy of economics (cf. Section 5.2) but, as I argue ahead, is only true within a model world and not when models of mechanisms are taken as evidence for policymaking.

There are a few obstacles to employing the mechanistic account of causality. First, Phil Dowe (2001) argued that the mechanistic approach does not deliver a tool for analyzing causation by omission. This problem is irrelevant for economics since, in line with Dowe's conclusion, economists usually employ the counterfactual account for singular causal claims (cf. Chapter 3). Second, Williamson (2011b) indicated that an intuitive understanding of causality puts, according to recent psychological studies, emphasis not only on mechanisms, but also on difference-making. Given that many mechanisms operate at the same time, the knowledge of one mechanism may be insufficient to predict the effects of an intervention. Third, some mechanistic theories can be accused of circularity: they define causal relations as a result of mechanisms and mechanisms as parts interacting causally (Williamson 2011a). Due to the novelty of the mechanistic approach, there are many unresolved issues. For instance, should mechanisms be required to be stable? Do they have to produce the phenomenon regularly, or can a single causal chain also be labeled in this way? What are the parts that constitute mechanisms? How should causal relations between these parts be defined? (cf. Craver and Tabery 2017).

One of the questions regarding the ontology of mechanisms is whether they are something different from causal laws. Glennan (2002, p. 344) disagreed. According to his viewpoint, causal laws underpin mechanisms; entities that are parts of a mechanism interact with each other according to causal laws and generate a mechanism. However, the mainstream stance is the opposite: while causal laws can be either interpreted agnostically or in a reductionist way regarding phenomena underlying them, most mechanistic theories of causality are a realist about what produces causal relations. An exception is Beck's (2006) view that it is impossible to 'observe' causal processes and mechanism because they operate on the lower level of reality and therefore are unobservable. Interestingly, Cory Wright (2012) argued that discussing 'mechanistic explanation' is elliptical because the notion of 'mechanism' is inherent in 'explanation.' On this basis, Wright criticized an explicit commitment to the ontological dimension of mechanisms: explaining presupposes that the postulated mechanisms are true because only true accounts explain (Reiss 2012). The mechanistic approach is also criticized for its lack of strictness and limiting the aim of science. Tracing mechanisms may not be suitable for social science. Reiss (2007, p. 168) listed the following four aims of social scientists: theoretical explanation, description, prediction, and control, and argued that mechanistic analyses focus on the first goal. On this basis, he disagreed with the critical-realist emphasis of the 'new mechanist perspective' that diminishes the importance of other aims of science. Limiting the goal of science exclusively to explanation is connected to the critical-realist presuppositions on the realm of social sciences.

5.1.3 Mechanisms in the philosophy of economics

Despite criticizing mechanistic philosophy, Reiss (2007) admitted that mechanisms have recently become a central unit of analysis in the social sciences. Uskali Mäki (2009a) acknowledged that 'mechanism' "is one of economists' favorite words used in a variety of contexts such as kinds of market mechanism, incentive mechanism, and transmission mechanism" (p. 85). Caterina Marchionni (2017, p. 434) underlined that "[k]nowledge of mechanisms has been claimed to play a key role in making causal inferences from statistical data more secure." In this section, I review the discussions of mechanisms in the philosophy of economics. First, I address the question of what economic mechanisms (i.e., mechanisms within economics) are. Second, I consider the critical realist distinction between closed and open systems that proves useful in developing my argument that mechanistic knowledge is insufficient for predicting the outcomes of interventions in Section 5.3. Third, I discuss the mainstream view that theoretical models represent (sometimes simplified) mechanisms isolated from the influences of other factors).

Beforehand, I need to introduce the notion of a theoretical model. These, contrary to the econometric models discussed previously, are deductive structures. To put forward a theoretical model, economists start from delivering a set of propositions (axioms). A solution can be mathematically deduced (which usually involves solving differential equations) from these axioms. A paradigmatic example of how theoretical modeling is used to isolate one causal mechanism is Johann von Thünen's model put forward in his 1863 book known as *The Isolated State*[2] (Thünen et al. 2009). The model, which is a case study often used by the philosophers of economics, is devoted to studying the economical use of land. In the simplest version of the model, Von Thünen assumed that the spatial differentiation of land use is determined exclusively by transportation cost. The model entailing several idealizing assumptions (such as the assumption that there are no roads in the 'isolated state' and the cost of transportation depends exclusively on distance) describes the mechanism of locating production of dairy and market gardening, wood, and cropland. What follows from the assumptions chosen by von Thünen is that these three different commodities are located in circles. However, considering that there are no countries in the world that use land in a way producing the circular pattern, the question arises of how such highly abstract models relate to the world.

Mäki (2008, p. 54) described the process of constructing such models in terms of experimentation:

> [i]n analogy to experimental procedure, such idealizing assumptions in many contexts serve the further strategic purpose of theoretical isolation. By neutralizing other subsidiary causes and conditions, they help isolate a major cause and its characteristic way of operation. . . . What is isolated by his simple model is distance (or transportation costs) as a major cause of land use distribution.

However, economic models of phenomena can also be considered (in some cases, at least) as pieces of mechanistic explanations (Gerring 2010, p. 1505). These mechanistic explanations are also partial. For example, Petri Ylikoski (2012, p. 24) admitted that a "mechanism-based explanation describes the causal process selectively. It does not aim at an exhaustive account of all details but seeks to capture the crucial elements of the process by abstracting away the irrelevant details." In other words, economists isolate irrelevant factors to get a model of the mechanism under consideration in a way similar to the process of experimentation in the natural sciences, where scientists create artificial conditions to test the influence of one factor. For instance, testing Newton's law of gravity requires creating a vacuum.

Are economic mechanisms distinct from those operating in the realm depicted by other sciences? Caterina Marchionni (2017, p. 427), who recently analyzed the meaning of mechanisms within economics, defined them as "complexes of rational agents, usually classified into social categories, whose actions and interactions generate causal relationships between aggregate-level variables." While such a definition rightly grasps the models represented by main neoclassical mechanisms, I need to emphasize that some contemporary advanced models deviate from the assumption of rationality, and therefore they stand for mechanisms populated by both rational and irrational agents.

Two distinctive understandings of mechanisms were depicted previously. On the one hand, mechanisms can be understood in the realist way; i.e., accepting the existence of entities and interactions between them that produce a phenomenon under consideration. On the other hand, mechanisms can be identified with explanations. In this case, they are identified with pieces of theories that explain (i.e., deliver a description of lower-level entities and interactions between them) with disregard for their 'real' existence. Accepting such a definition is clearly antirealist in nature. Marchionni highlighted that her definition is descriptive (i.e., not normative) in nature and results from analyzing what economists take as mechanisms. In other words, her goal was to reconstruct the meaning of 'mechanism' by studying how economists represent them. Therefore, assuming the fulfillment of the undertaken goal, economists are realists about mechanisms. Furthermore, Marchionni (2017, p. 427) indicated two distinctive features of how mechanisms are represented. First, economic mechanisms are described by assumptions grounded in the rational-choice theory. Second, economists make use of highly abstract axiomatic models that hugely rely on mathematics.

Julian Reiss (2013, pp. 104–105) identified four distinct notions of 'mechanism' present within economics. First, econometricians working in the Cowles Commission tradition label single equations of the structural-equation models 'mechanisms' referring to their causal but not necessarily mechanistic nature as understood in this chapter. Second. 'mechanisms' can refer to mediating variables. For instance, if $A \rightarrow B \rightarrow C$ (i.e., B is the intermediary variable), then A can be said to cause C by the mechanism of B. Third, mechanism can be understood as a structure (system) or process that generates a phenomenon under consideration. In this case, 'mechanism' refers to "processes that lie at a deeper level" (p. 105). Finally, economists label in this way fragments of a theory that describe a part of the world. Of these four different meanings, only the latter two understandings overlap with how mechanisms are defined in the philosophy of causality discourse. The 'mechanism as underlying structure or process' notion is in line with the realist understanding of mechanism as entities interacting with each other and producing phenomena. The 'mechanism as a piece of theory' idea is in line with the antirealist understanding of the philosophy of causality.

Contrary to Julian Reiss, Marchionni (2017, p. 424) defined mechanisms in line with the 'new mechanists' movement in the philosophy of causality; i.e., in terms of "underlying structures [because the meaning is . . .] closest to the conception of mechanisms advanced by current mechanistic philosophers." She further developed her understanding by employing Illari and Williamson's (2012, p. 120) definition: "[a] mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon." Accepting such a definition highlights the ontological understanding of 'mechanism.' Knowledge about mechanisms connecting two variables is generally believed to confirm causal inferences put forth on the grounds of statistical methods (Marchionni 2017; Steel 2004). In fact, Jon Elster (1983) argued for a skeptical viewpoint, according to which drawing causal conclusions from observational studies is only justified if an investigator possesses knowledge about mechanisms. However, accepting this stance presupposes that mechanistic knowledge is less prone to error (not to say that it presupposes its certitude). The presupposition is problematic especially in the social sciences, where testing whether a represented mechanism is right is difficult or impossible.

Alex Broadbent (2011, p. 59) argued that the need to infer mechanisms arises when a previously unknown dependence is established and we seek an explanation for it: "[w]e make warranted inferences to causal generalisations; these generalisations imply the existence of underlying mechanisms; and we then conduct further research to find the mechanisms." How the causal inferences employing the notion of mechanisms proceed? According to Julian Reiss (2013, p. 105), "to provide a mechanism for an aggregate relation, then, is to describe how the entities and processes that underlie the aggregate variables are organized and interact with one another in such a way that the aggregate relation results." Inferring mechanisms is an epistemically demanding task that may fail policymakers' expectations. This is the case due to the fact that causal mechanisms "do not necessitate their effects" (Reiss 2013, p. 106). Reiss (2013) exemplified his viewpoint with the case of channels of monetary transmission discussed by Mishkin (1996). The presence of several 'channels' or mechanisms creates a situation whereby "it is not normally possible to predict an outcome upon observing that this or that mechanism has been triggered" (p. 108).

Similarly, I argue ahead that the knowledge of one mechanism is not sufficient for predicting the outcomes of policy interventions. The distinction put forth by critical realists between open and closed systems (cf. Bhaskar 2008, ch. 2) sheds light on why this is the case. In closed systems, like an experimental setting (e.g., bodies falling in a vacuum) or a clock, one or a few mechanisms operate without external influences. However, in open systems, several mechanisms operate at the same time, and therefore one mechanism may multiply the effects produced by another one, hamper them, or interfere in still different ways.

What follows is that if there are many mechanisms operating at the same time and we do not know how they interact with each other, then even accurate knowledge of one mechanism is insufficient for predicting how the whole system behaves. Given that the mainstream view on theoretical, axiomatic models is that they represent only one isolated or/and idealized mechanism, having one true model is insufficient for prediction of the effects of interventions without knowledge of how different mechanisms interact with each other.

Mechanisms in social reality

Scientific realism and critical realism are two realist stances in the philosophy of economics that differ regarding the views on mechanisms. However, mechanisms, nevertheless, are their central tenets. Scientific realism reconstructs the views of economists on epistemology and ontology by studying the research practice of economists. On the contrary, critical realism belongs to the branch of philosophy known as social ontology and is normative in nature. Tony Lawson (1997), the most notable critical-realist economic methodologist started his research from presupposing certain features of social ontology and produces epistemic advice on this basis.

Mechanistic analyses are grounded in demand for microfoundations that results from the ontological presupposition that the reality (of social sciences) is stratified. In detail, the investigation of causal mechanisms is explicitly advised by Lawson (1997), who employed the critical-realist stance that there are different layers of reality. Roy Bhaskar (2008, p. 47) listed the level of (1) experiences, (2) events, and (3) mechanisms. The operation of different mechanisms and interactions among the three levels of reality undermine causal inferences in the realm of social sciences. As he put it,

> different levels . . . mesh together in the generation of an event . . . and [events] will not normally be typologically locatable within the structures of a single theory. In general the normic statements of several distinct sciences, speaking perhaps of radically different kinds of generative mechanism, may be involved in the explanation of [one] event.
>
> (Bhaskar 2008, p. 115)

According to the critical-realist stratification, causal mechanisms underlie the level of events and are responsible for their production. While critical realism and new mechanistic philosophy are different in several aspects, both stances accept that the reality is stratified. For instance, one of the mechanistic philosophers, Glennan (2012, p. 317), indicated that mechanisms are hierarchical what makes the new mechanistic philosophy in agreement with critical realism. However, not every kind of mechanism is in agreement with the social ontology put forth by Bhaskar (2008) that inspired the world view of Tony Lawson (1997). Reductionist mechanisms that contradict the critical-realist ontology are criticized on this basis. For example, Geoffrey Hodgson (2004) focused on the use of 'Darwinian' mechanisms by evolutionary economics. Such mechanisms use the options of survival, selection, and reproduction of structures to explain social phenomena. Hodgson (2004, p. 185) defined mechanism as "a structure involving causal connections but lacking an adequate capacity for self-reflection, intentionality or will," and argued that the use of Darwinian mechanisms in economics is grounded in the presupposition that even human decisions are caused by certain factors what undermines human free will.

The projects of critical realism and scientific realism (described ahead) are divergent regarding the possibility of isolating causal mechanisms by means of theoretical modeling. Lawson (1997, p. 31) highlighted that the mechanisms operating in the social sciences could not be isolated due to the inherent openness of the social world. Accepting that different causal mechanisms operate at the same time at different layers of reality leads to the conclusion that prediction and control are impossible. Even understanding the work of all the mechanisms operating in a particular context is not sufficient: fulfilling the aim of prediction and control demands the need of understanding how the influence of different mechanisms can be summarized. Mervyn Hartwig (2015, p. 231) added still another obstacle on the path to accurate inferences of mechanisms: social mechanisms, according to critical realists, evolve and therefore they are not universal (they are ephemeral).

Theoretical models as representations of mechanisms

An example of mechanistic modeling in economics is the Schelling's (1969) widely discussed checkerboard model (Aydinonat 2007; Grüne-Yanoff 2009).

The model is usually said to explain[3] empirical regularity (spatial segregation of races in American cities) with a lower-level mechanism (black and white people, and their preferences of living in same-ethnicity neighborhoods) (cf. Ylikoski and Aydinonat 2014; Sugden 2009; Hardt 2014). Othertimes, Schelling's educational model is accused of being construction of little resemblance to the world (Hardt 2017, pp. 133–168). The model serves well as an example of isolation and/or idealization of one mechanism (Mäki 2009a).

Thomas Schelling started his modeling exercise with several assumptions describing the following situation: there are pluses and minuses randomly located along an axis:[4] the signs prefer to have neighbors of the same kind. At each stage, pluses and minuses can exchange their position if they are neighbors. The model shows that even a slight difference in preferences can lead to strong segregation after several iterations. Of course, as every model in economics, the checkerboard model employs mathematics and is formalized. Considering that the location one lives at is determined not only by one's preferences but also several other determinants such as the prices on the housing market and one's earnings, Schelling's model does not represent the phenomenon fully but stands for only one causal mechanism among many at work Therefore, it is a partial representation. In other words, some factors are isolated away.

Furthermore, it may be the case that the actual mechanism that is responsible for the creation of racial segregation is different (e.g., the differences in incomes and the spatial dispersion on the housing market) from the one represented by Schelling's model. Therefore, the model represents a possible mechanism (a mechanism that could produce the phenomenon under consideration) and not the actual mechanism that indeed produced it. Based on this, N. Emrah Aydinonat (2007 p. 430) labeled the checkerboard model as "a partial potential theoretical explanation in the sense that it suggests some of the mechanisms that may bring about residential segregation."

The model exemplifies how economists infer causal mechanisms by means theoretical modeling. It can be divided into three stages. First, a phenomenon seeking a mechanistic explanation is identified. The candidates for such phenomena are usually those relations that are suspected of being causal. For instance, an econometric research can suggest that there is a dependency between two variables or the everyday experience delivers inspiration for a causal-mechanistic research (cf. Krugman 1993). Second, axioms that describe a postulated mechanism are put forward. Finally, a model of mechanism should be appraised regarding its resemblance by the community of economists. However, as I argue ahead, the last step is quite problematic.

One of the main roles of theoretical models is to deliver explanations for observed outcomes. The mechanistic explanations (which are grounded in the presuppositions that the employed models are models of mechanisms) are present in economics since the beginning of the discipline. Adam Smith's (1950 [1776]) discussion of how the invisible hand coordinates socioeconomic decisions is, despite the metaphorical language, a mechanistic analysis (cf. Ullmann-Margalit 1978). As Marchionni (2017, p. 423) put it: "[t]he market represents the paradigmatic example of an economic mechanism. Adam Smith famously theorized it as if led by an invisible hand to satisfy the needs of market participants." Mechanisms also play a crucial role in the works of Uskali Mäki (2013, 2011a, 2011b, 2009a, 2008, 2005), who argued that:

> [b]y representing a mechanism inside an input-output system, economists not only convey knowledge that the input and the output are connected, they also conjecture how the input, together with the mechanism, produces the output. . . . And answering such *how* questions enables economists also to be more assured *that* there is a causal connection between *I* and *O*, thereby establishing a causal relationship where there appeared to be mere correlation or empirical regularity.
>
> (Mäki 2009a, p. 86)

From studying how economists construct their theoretical models, one can learn that the models are isolations and/or idealizations of certain aspects of the economic world (Mäki 2005; Morgan and Knuuttila 2012). Economists explain by delivering models that isolate and idealize causal mechanisms (Mäki 2009). Uskali Mäki and Bruce Caldwell (1992) argued that the purpose of economic models is to isolate (and, in some instances, idealize) 'causal processes.' Scientific realists (at least in general philosophy) accept the knowledge thesis: our most developed current theories are, at least approximately, right about the world (Papineau 1996, p. 2). In the philosophy of economics, the scientific-realist project seems to be less optimistic, especially after the 2007–2008 financial crisis that shed light on the inadequacy of certain assumptions (e.g., rationality) that are crucial, especially for the mainstream neoclassical modeling (Krugman 2009; Hardt 2016). It is interesting to observe that, over time, Mäki's views evolved so that his model of modeling (ModRep) now includes several pragmatist dimensions and the refusal of the idea that the theoretical economic models rightly grasp the reality. Reiss (2013) doubted if scientific realism (in its philosophy of economics version) would be classified as such a doctrine due to a substantial overlap with pragmatism. Nevertheless, Mäki (2009b, p. 92) defended the unrealism (i.e., falsity) of economic models, indicating that the purpose of the models of phenomena (cf. Frigg and Hartmann 2006) differs from being adequate to data: "[t]his purpose is that of theoretically isolating some important dependency relation or causal factor or mechanism from the involvement and influence of the rest of the universe" (Mäki 2009b, p. 92). However, he later added that the realist commitment demands from such models to appropriately describe real mechanisms: "the mechanisms in operation in those imaginary situations are the same as, or similar to, those in operation in real situations" (p. 95). To put it in other words, "[a] model captures significant truth if it contains a mechanism that is also operative in real systems" (ibid.). Unrealistic (i.e., literally false) assumptions are acceptable with the aim of isolating these mechanisms from the influence of external factors.

A widely felt intuition states that only actual mechanisms explain. That is, theoretical models must describe mechanisms that operate in the world. However, there are severe doubts about the veracity of economic models (e.g., Hausman 1992, p. 121). The problem with the models of mechanisms in economics is that the theoretical models are often considered to be literally false. False models cannot explain because they deliver descriptions of nonexistent mechanisms (e.g., Mäki 2006). Based on this, Reiss (2012) formulated the following explanation paradox (known as Reiss' trilemma): "(1) economic models are false; (2) economic models are . . . explanatory; and (3) only true accounts explain" (p. 43). Each pair of the three statements is plausible, but they are contradictory if taken together. Several philosophers have an issue with each one of the three theses. Anna Alexandrova and Robert Northcott (2013) opposed Reiss' second commitment and put forth the view that economic models do not explain. Only models of actual mechanisms can explain.

Mäki (2009a) defended the project of mainstream neoclassical modeling by interpreting models as 'credible surrogate systems' that are studied instead of researching the economy due to its epistemic inaccessibility. Such surrogate models are highly abstract and isolating. They "are often about single mechanisms" (Mäki 2009a, p. 29). Modeling a single isolated mechanism makes it possible to analyze how the mechanism under consideration works. The isolation "involves . . . control for noise so as to isolate some important fact, dependency relation, causal factor or mechanism" (Mäki 2009a, p. 30). An example of such a highly abstract piece of economic modeling is Schelling's (1978) checkerboard model interpreted as a model of mechanism how singular preferences of city inhabitants determine macrobehavior (Sugden 2000; Aydinonat 2007). Mäki (2009c, p. 10) reads economic models as "imagined systems in which a simple streamlined mechanism is in operation isolated from any other complexities and interferences."

In *Economics Rules*, Dani Rodrik (2015) presented a similar view of the theoretical models in economics. According to his account, theoretical economics is a pluralist discipline in the sense that economists deliver many models for the same phenomena. The models that are useful for a given purpose (prediction, explanation, etc.) are chosen from *a library of models*. The library includes models that represent both actual and possible (fictitious) mechanisms. Accepting this leads to the view that it is not the task of a modeler to establish that their model represents the mechanism that produced the phenomenon seeking an explanation; it is the model users' responsibility. Rodrik, unfortunately, described his views on how models are chosen quite vaguely. Till Grüne-Yanoff and Marchionni (2018) attempted to formalize his account. N. Emrah Aydinonat (2018) argued that multiple models can be used to improve the adequacy of economic explanations. However, still considerably little attention has been put on addressing the question of how right and wrong models of mechanisms are discriminated.

The view that all axiomatic models represent mechanisms has been contested in the literature on the grounds that such models can serve different purposes and explain in different ways. Hardt (2017) pointed out that economic models deliver causal explanations (that can but do not have to be mechanistic), mathematical explanations, and statistical explanations. In the case of mathematical explanations, there is no need to employ causes. The question if mechanistic explanations are crucial for economists stays open. For instance, Verreault-Julien (2017) considered the case of the general-equilibrium model (which is the most notable example of the mainstream neoclassical modeling tradition) and argued that causal explanations are unjustified in this case.

Finally, I need to introduce the distinction between causal (and non-mechanistic) and mechanistic explanation. On the one hand, causal (non-mechanistic) models represent regularities found in the world or in an aspect of the word by means of mathematical, functional dependencies. An example of a purely causal model is the theoretical law of demand that describes the influence of a change in the price of a good p on the quantity sold on a market q. Taking a textbook example from physics, the law of gravity can be interpreted causally in this sense. Given that such laws (being models) do not describe elements of the system and interactions among these elements, postulating that such models represent mechanisms is not warranted. Such models do not explain causal relation by delivering a mechanism producing it but limit its purpose to depicting a (multidimensional) law of nature using quantitative functions. To the contrary, mechanistic models represent mechanisms: they describe lower-level entities and their interactions that produce a phenomenon (regularity) at a higher level of reality. Most contemporary theoretical models published in mainstream economics journals seem to represent mechanisms.

5.2 Developing mechanistic models

Economists use theoretical, axiomatic models to infer the mechanisms that operate in the economy. In this section, I analyze three examples of theoretical, axiomatic models that are representative of different modeling practices present in the mainstream. Highly abstract, axiomatic models (such as Schelling's checkerboard model considered previously) are purely theoretical (nonempirical); i.e., do not use data in any formalized way. The whole input to such models are axioms chosen by modelers. On the contrary, calibrated models base the values of some parameters on 'calibration.' As Christina Dawkins (2001, p. 3656) observed, "[c]alibration . . . remains an imprecise term despite its widespread use." Two types of the process can be distinguished.

On the one hand, calibration can denote either choosing the values of parameters on the basis of values delivered by statistical agencies (e.g., last year's value of inflation) or using econometric techniques to 'fit' a model into data of interest. DSGE models are a distinct subgroup of calibrated models. The purpose of these structural dynamic programming models is to represent the whole economy. They are micro-based macroeconomic models that entail different groups of entities (such as companies and customers). Heterogeneous agents are present in cutting-edge models after Iacoviello's (2005) move toward this more realistic assumption.[5] DSGE models are becoming more and more popular in recent years. The change in economic modeling is probably driven by the technological revolution allowing for computational tractability of these highly advanced models and the wave of skepticism against highly abstract axiomatic models after the 2007–2008 financial crisis.

This section is structured as follows. First, I consider Matthew Kotchen's (2006) game-theoretic model of a market mechanism for green goods having features of both public and private goods (4.2.1). Second, I discuss Ian Parry and Kenneth Small's (2009) calibrated model of a mechanism underlying the relationship between public transport subsidies and transport efficiency in agglomerations. Third, I study Lawrence Christiano et al.'s (2011) DSGE model devoted to studying government spending multiplier. Finally, I discuss the use of case studies for inferring mechanisms. Despite its limited use in mainstream economics, the qualitative approach deserves some attention due to its popularity in related disciplines (Section 5.2.4). The following section (5.3) discusses the use of theoretical models for policymaking.

5.2.1 A purely theoretical model

An excellent example of an axiomatic model that postulates a possible mechanism is Kotchen's (2006) game-theoretic model of market for impure public goods.[6] The causal claim supported by Kotchen's model is that "changes in equilibrium provision will depend on preferences, the distribution of income, and the green technology." (p. 826) The model is a model of mechanism (it fulfills Marchionni's (2017) definition of economic mechanism), given that the causal relation depicted by the claim cited above is produced by the interactions of entities (consumers with their preferences and green goods of specific characteristics). Interestingly, Kotchen (2006) uses the word 'mechanism' four times in his paper. Three of the uses are in line with the philosophical meaning of the term. For instance, at the end of the paper, one can find the assertion that "[t]he results . . . provide new insight into the potential advantages and disadvantages of green markets as a decentralized *mechanism* of environmental policy" (p. 831) (italics added). One time, 'mechanism' denotes a quantitative technique allowing for reproducing the results of a game, including a small number of participants into a bigger economy (p. 826).

Even though Kotchen started his paper by delivering statistics in favor of the growth of environmentally friendly goods and services, his modeling exercise is not influenced by data in any formal or systematic way. Given this, his model is representative of highly abstract deductive models. The preceding review of the philosophical views on theoretical modeling in economics indicates that relatively little has been said on how mechanisms are inferred with theoretical models. Considering that the views that theoretical models possibly only deliver how-possibly[7] explanations (Mäki 2013; Verreault-Julien 2019), and the role of theoretical economists are to deliver many different possible explanations (Rodrik 2015), the question of whether Kotchen's (2006) model represents an actual or a possible mechanism stays open. Another exciting issue that received limited attention is the process of choosing[8] axioms on which theoretical models rely. Leaving this topic beyond the scope of philosophical investigation of modeling can be explained by the fact that, before the era of the social studies of science began in the late twentieth century (e.g., Latour and Woolgar 2013), the context of discovery (i.e., how scientists get the idea for their theory) has been excluded from the endeavor of philosophy of science. While my book by no means attempts at belonging to the sociology of science, considering how theoretical economists choose the assumptions of their models (given that these assumptions deductively determine the results) and whether they support the claim that the within-model mechanisms actually operate in the world is of interest for the audience of economics research. Kotchen's (2006) paper can be used to study these questions because, as the author admitted, the text addresses the question "about how the option to consume impure public goods affects private provision and social welfare" (p. 817). That is, the paper delivers an explanation.

The axioms put forward by Kotchen (2006, pp. 819–820) describe the modeled world populated by *n* consumers that purchase goods having two characteristics: X (the properties of a private good) and Y (the properties of a public good). These individuals earn exogenous wealth $w_i > 0$ and allocate it among the following three types of goods:

1 Pure private goods having only the characteristic X.
2 Pure public goods having only the characteristic Y.
3 The 'green' goods having a mixed characteristic of X and Y.

Each consumer chooses goods in order to maximize their utility function depending on the consumed amount of private goods X_i, financed public goods Y_i, and public goods financed by all other consumers Y_{-i}. The maximization problem is limited by the budget constraint, and X_i, and Y_i, are constrained by the number of goods purchased by consumer *i* and the society. In addition to assumptions describing the agents living in the model world, Kotchen (2006) needs to construct the supply side of the economy. The three types of goods are sold at the prices equal to one for the amounts of c_i, d_i, g_i. These goods are normal goods (i.e., the elasticity of demand for each good is positive and lower than 1). Each type of good delivers a certain amount of characteristics X_i and Y_i. The amounts are specified as follows. One unit of good c_i, delivers one 'util'[9] of X_i and one unit of good d_i delivers one unit of 'util' Y_i. That is, *c* and *d* are goods (products), while X and Y denote utility (pleasure) delivered by these goods. Kotchen (2006) also assumes that the green good g_i is more 'efficient' in delivering both X_i and Y_i in comparison to purchasing a combination of goods c_i, and d_i. This assumption (labeled henceforth 'technical assumption') proves problematic (more on this ahead).

Before discussing the results that follow from these assumptions, I need to consider the nature of the assumptions chosen by the economist. While some simplifying assumptions can be of technical nature that allows for mathematical tractability of a model, others isolate the postulated mechanism from external influences. Still others distort the modeled mechanism by modifying it in crucial ways. The assumption that the wealth w_i earned (owned) by consumer i is exogenous seems to exemplify isolating assumptions. Despite being literally false (one's wealth, at least partially, depends on previous decisions, efforts, etc., and therefore is not exogenous to agents' decisions), it does not distort the modeled mechanism: extending the model in a way that would include a mechanism determining consumers' wealth would not change decisions regarding the consumption of green goods. Another example of unrealistic assumption that does not distort the modeled mechanism in a meaningful way is setting the prices of goods to unity and letting the amounts of goods vary. Since our common sense allows us to conclude that the prices are set in euros per a specified amount (e.g., one piece, kg, liter, etc.), it is descriptively false. However, despite diverging from the standard practice, a simple mathematical transformation allows for obtaining prices in the standard notation. Therefore, this assumption simplifies mathematical tractability of the model without influencing the represented mechanism in a significant way.

However, the character of the technological assumption that the green good delivers both X and Y at a lower cost than purely private or purely public goods is very different (Kotchen 2006, p. 820). Even though this assumption also allows for mathematical tractability of the model by reducing the number of equilibrium states to one, it also, if false, distorts it in a severe way. Considering that Kotchen (2006, p. 827) later considered the effect of assuming the opposite of the results, his modeling endeavor confirms the view that economic modeling is a robustness analysis (cf. Kuorikoski et al. 2010). Unfortunately, Kotchen (2006) did not support this assumption in any way, even though it is crucial for the results of the modeling exercise. While the technological assumption may be plausible in some cases, it may also be false in others. For example, it may be true that buying 'shade coffee' is more cost-effective than buying both usual coffee (private good) and donating for a rainforest-oriented foundation (e.g., Greenberg et al. 1997). However, the assumption may be false regarding other goods. For example, considering the number of external effects of the process of producing batteries to electric cars (e.g., Racz et al. 2015), buying a much cheaper traditional car and offsetting CO_2 emission by donating to a tree-planting charity may be more efficient. If the technological assumption is modified accordingly (i.e., stating that buying purely private good and purely public good is more cost-effective than purchasing the green good), then the result of the model is overturned: it is not buying the green good and either private or public good (depending on preferences) that is the most cost-effective choice, but a combination of public and private goods.[10]

After putting forward the assumptions previously discussed, Kotchen (2006, p. 823 et seq.) discusses the different scenarios (propositions) specifying the green goods. This shows that the understandings of causality are interrelated. Given the different features of green goods introduced into the model economy, the presupposed meaning of causality possibly overlaps with the counterfactual approach. However, considering that (1) the main conclusion that the equilibrium state obtained after the introduction of a green good depends on "preferences, the distribution of income, and green technology" (Kotchen 2006, p. 826) which can be also derived deductively from the set of assumptions (what Kotchen does in the appendix, pp. 831–833), and (2) the model represents a mechanism, the mechanistic view on causality seems to dominate the presupposed meaning. The causal claim is of type-level nature that additionally differentiates this claim. What are the relata?

The features of aggregate phenomena seem to be the right candidate. First, while the model describes a mechanism (i.e., the interactions of agents facing a market for public and private goods), the causal claim describes the causes and effects of an aggregated phenomena. In other words, the causal relationship between these phenomena is produced by (emerges from) the decisions of agents. Second, equilibrium states, (features of) green technology, and distribution of income are definitively features of market exchange, production of green goods, and the effects of social interactions, respectively.

Addressing the question of whether Kotchen's (2006) model represents an actual or a possible mechanism (i.e., whether it operates in reality) is crucial to use mechanistic evidence for policy purposes. Unfortunately, the economist has not delivered empirical evidence in support of the actuality claim. Therefore, while the model can be taken as a representation of a possible mechanism (that could produce the observed result), it is epistemically unjustified to accept it as a representation of an actual mechanism. As Verreault-Julien (2019, p. 32) admitted, "[i]n some cases, we may lack the empirical support to establish a claim of actuality, but may have enough for a possibility claim." In the case of the discussed highly abstract theoretical model, the empirical support is missing.

5.2.2 Calibrated theoretical models

Calibrated theoretical models are similar to the type of theoretical models previously discussed in being axiomatic, deductive systems. However, contrary to the previous group, calibrated models are based not only on assumptions, but also have empirical input. A representative example of this type of evidence is the model constructed by Ian Parry and Kenneth Small (2009) with a view to mechanistically analyze the influence of public transport subsidies on welfare. Addressing the research question is motivated by the fact that public transport is heavily subsidized in most developed countries, and while purely theoretical modeling (e.g., Glaister 1974; Henderson 1977; Jansson 1979) delivers argument for such a policy, the question regarding the efficient level of subsidies remains to be answered on empirical grounds. As the authors admitted, previous studies have focused on only one city and a limited number of factors what resulted in obtaining "estimates of optimal transit prices [that] vary enormously . . . providing a confusing guide as to whether current fare subsidies should be preserved" (Parry and Small 2009, p. 701).

Despite the empirical input of Parry and Small's (2009) model coming from the calibration process, the model is similar to the theoretical model in isolating away unimportant factors from the represented mechanism. The rationale for excluding some factors is the striving for the generalizability of the results and quantitative simplicity (comparing to studies incorporating a broader set of factors). As Parry and Small (2009, p. 703) put it, "models . . . describing individual users with heterogeneous characteristics, on the demand side . . . are difficult to compare across regions and cannot easily provide transparent intuition about the underlying reasons for particular results." However, despite some simplifying assumptions, calibrated models are usually sufficiently advanced (in terms of mathematical structure) to grasp complicated mechanisms. I need to highlight that the axiomatic models of this type are inferior in prediction compared to simpler econometric models. However, their advantage is the interpretability of the results. According to James Heckman (2000, p. 50), "[c]alibrators emphasize the fragility of macro data and willingly embrace the conditional nature of causal knowledge. They explicitly reject 'fit' as a primary goal of empirical economic models and emphasize interpretability over fit." The calibrated models are superior to purely theoretical models in empirical fit and inferior to econometric models in prediction accuracy, but beat them regarding understanding of represented phenomena.

The modeling practice starts from constructing mathematical structure representing mechanism. Similar to how purely theoretical modeling proceeds, the task is achieved by putting forward a set of axioms. One of the simplifying assumptions that is commonly employed in micro-based macroeconomic models is the use of a representative agent instead of modeling the divergent decision-making processes of a diverse population. Such an assumption precludes the modeler from posing questions regarding disaggregated phenomena. For example, it is plausible to suspect that commuters' decisions are heavily influenced by their income. Using the representative (i.e., average) agent framework makes studying the effects of income distribution impossible. Deciding whether such assumption distorts the aggregate-level conclusions requires possessing knowledge of effects of income distribution.

On the one hand, the use of representative agent can only simplify the model by isolating the represented mechanism from the effects of income differentiation. On the other, it may distort the mechanism in a serious way. Parry and Small (2009) believe that "considering separate income groups would add flexibility important for certain questions such as the effects of differentiated products" (pp. 703–704), but excluding it from the model does not distort the results given that the model includes the calibration for demand elasticity that results from income differentiation. Deciding on which of the two possibilities is the case in this situation likely requires specifying beforehand the causal questions addressed by the model. At the end of the paper (p. 722), the economists admit that the passengers using buses have usually lower incomes comparing to the customers of rail transit, but it does not influence the conclusion regarding the aggregate-level subsidy level.

There are further simplifying assumptions.[11] All features of transportation (congestion, travel time, transit frequency, vehicle crowding, etc.) are summarized into a variable denoting 'cost' that is considered by agents in their decision-making process. The length of trips is assumed to be constant, so the variation in traveled distance is included as a variation in the number of trips. The representative agent maximizes utility choosing time and mean of transportation, minimizing the cost of travel, and maximizing consumption of other goods (numéraire)[12] and travel for all purposes (p. 704). The utility functions are quasi-concave, what implies that the time of travel and different models of transportation are imperfect substitutes. The disutility from pollution and traffic accidents is included in the model. After putting forward assumptions that specify the optimization problem of the representative agent, Parry and Small (2009, pp. 706–708) describe axiomatically the modes of transportation (e.g., rail transit, car, bus, etc.) and their features (per passenger cost of one mile, vehicle occupancy, etc.). Later, they put forward the optimization problem of 'transit agency' responsible for setting public transport prices and financing its deficit from taxes.

After specifying the mathematical structure of the model by putting forward assumptions and describing the decision problem, Parry and Small (2009) calibrated their model for London, Washington, and Los Angeles. Given the lack of needed information, Parry and Small (2009, Appendix B) gathered extensive data from different sources (such as statistical offices, public transport companies, and environmental agencies) and estimated all the necessary values for the three different cities, means of transportation, and peak and off-peak travel times. Contrary to the purely axiomatic models that are usually based on general assumptions (e.g., that the production of green technology is more efficient; $X > Y$ in the case of the model of green goods market considered previously), calibrated models include numerical values for the variables. For instance, the cost of travel time includes in-vehicle time and wait cost, and they differ substantially for different cities and times of day. In Washington, in-vehicle travel time per passenger mile costs 73 cents in peak time and 47 cents in off-peak time. However, the wait time is 87 cents in the peak time and 38 cents in the off-peak time. The difference makes the overall travel time 'cheaper' in the peak time.

While the economists exported the discussion of the calibration process from the paper to the Appendix B and even a brief review of each calibrated parameter would definitely bore even most patient readers, considering one case is useful for addressing the question of whether this model represents an actual or a possible mechanism. To estimate average wait time, Parry and Small (2009, p. 714) analyze the time between subsequent connections and assume the following time management rule used by the representative agent: "if vehicles are less than 15 minutes apart, travelers arrive at random . . . ; but as the time between vehicles exceeds 15 minutes, an increasing fraction of traveler use a timetable, thereby lowering the wait time." To calculate the monetary value of the wait time, the authors multiply the estimation obtained by the following procedure by a fraction of the market wage that is based on empirical studies of how public transport users praise waiting. The process shows that calibration (despite being empirically based) is not as straightforward as econometric modeling. On the contrary, the calibration of each parameter could possibly lead to obtaining a different result. For example, while the idea that the more often a bus/train goes, the less likely passengers are to check a timetable before arriving at a stop is plausible, other decision rules cannot be excluded. Assuming different elasticities of the demand for wait time may lead to different results.

To exclude the possibility that their result arises from very specific assumptions and values of parameters, Parry and Small (2009, pp. 718–719) conducted sensitivity analysis aimed at establishing that the result is robust to changes in assumptions and the values of calibrated parameters. They modified the values of parameters by 30% and 100% in both directions, and changed the assumptions describing the choice of subsidies by the public transport agency. The analysis showed that the main causal conclusion that raising transit subsidies from the current levels in London, Washington, and Los Angeles causes welfare gains is robust to the changes in model assumptions and parameter values.

This claim refers to the calibrated model of a mechanism of choosing the mode of transportation and transit subsidy levels. First, the model represents agents and their maximization problems. Second, the causal relation emerges from the interactions and decisions undertaken by the representative agent standing for commuters and the transit agency that decides on the public transport fares and subsidy level. Given that the model consists of rational agents and describes their maximization problems, while the causal claim at the aggregate level results from actions of and interactions between agents, the model fulfills Marchionni's definition of economic mechanism. The represented mechanism is isolated from the effects of some factors. A few assumptions idealize this mechanism by simplifying it (e.g., excluding the effects of income distribution on agents' decisions). Nevertheless, the model seems to represent an actual mechanism. If the difference between how-possibly and how-actually explanations lies in the empirical support for the latter (Verreault-Julien 2019), then this mechanism has obtained considerable amount of empirical support.

This support comes from the process of calibration. Purely theoretical models are based on assumptions without empirical input, and therefore interpreting them as models of actual instead of possible mechanisms requires additional evidence. Calibrated models are based on assumptions describing an actual economic system (i.e., agents, their preferences, and constraints), and therefore the interactions between them that produce the mechanism are likely to take place in the economic reality. Therefore, the model can be interpreted as representing an actual mechanism operating in London, Washington, and Chicago. The relata of the causal claim are 'welfare gains' and 'subsidy levels.' They seem to be aggregate-level variables that denote summarized features of phenomena (here: a change in overall welfare experienced by agents and prices of transit tickets).

5.2.3 The DSGE framework

Strictly speaking, dynamic stochastic general equilibrium (DSGE) models belong to the class of calibrated models exemplified previously with the model of public transport. However, they deserve my separate treatment for several reasons, but mainly due to their widespread use by central banks and governments as evidence for macroeconomic policy. As Kuorikoski and Lehtinen (2018, p. 3) admitted: "DSGE models are . . . built to capture new macroeconomic mechanisms." Such models fulfill Marchionni's (2017) definition of macroeconomic mechanism. The interactions among different types of macroeconomic agents (usually consumers, companies, and government) and their solutions to optimization problems produce higher-level (aggregate) causal relations. DSGE models are usually employed to predict how an economy will react to various shocks and interventions (Valdivia 2015). Their widespread use by central banks and governments for policy-related purposes makes DSGE models explicitly causal: their primary area of use is predicting the outcomes of monetary policymaking. Despite being usually outperformed by simpler econometric models in regard to predictive accuracy, DSGE models deliver interpretability and understanding macroeconomic mechanisms. Similar to other calibrated models, DSGE models utilize calibration to represent actual rather than possible mechanisms.

The contemporary DSGE models include heterogeneous agents. For example, household choices are modeled with differentiated utility functions with the aim of including the effects of wage differentiation on macroeconomic behavior (Fernández-Villaverde 2010, p. 26). In fact, the discussion of mechanisms is employed by the practitioners of DSGE modeling. For instance, the mechanism of wage rigidities is modeled with the Calvo-pricing rule that divides wages into its indexed and unindexed parts (Fernández-Villaverde 2010, p. 30) what is one of the ways in which prices influence real economy in the DSGE framework. The microeconomic behaviors are summarized to obtain estimates of the aggregate values. The topic of microfoundations was specifically addressed by Sbordone et al. (2010), who estimated a simple DSGE model for educational purposes. As the authors describe this aspect of their modeling exercise, the "model economy is populated by four classes of agents: a representative household, a representative final-good-producing firm (f-firm), a continuum of intermediate firms (i-firms) indexed by $i \in [0; 1]$, and a monetary authority" (p. 26). Each group of these economic actors acts according to their preferences so that the macrobehavior of modeled economy results from the microbehaviors of customers, producers, intermediaries, and the monetary-policy body. However, certain aspects of some DSGE models lack the microfoundational nature. For instance, the nominal rigidities of prices are added ad hoc (Fernández-Villaverde 2010). DSGE models are criticized for the lack of realism (e.g., Kirman 2010). The discussion of whether assumptions of economic models should be 'realistic' (i.e., descriptively accurate), and if unrealistic (i.e., false) assumptions can be allowed in some cases, is as long as economic modeling itself. In his famous essay, Milton Friedman (1953) allowed for the use of false assumptions as long as a model saves the appearances of things. For instance, the assumption of consumer rationality should be appraised if it allows for constructing a simpler model that scores better at prediction (cf. Maziarz 2018b). Others (cf. Section 5.1) disagree, and argue that unrealistic assumptions are only allowed if they aim at isolating or idealizing represented mechanism while modelers should refrain from the use of the assumptions that distort such mechanisms.

Another issue raised by the opponents of DSGE models is the question of whether the mathematical development (and its cost) pays off in terms of better predictive accuracy. Simpler econometric tools such as vector autoregressive models (VARs, cf. Chapter 3) outperform DSGE models (Del Negro and Schorfheide 2013; Smets and Wouters 2003). Nevertheless, this modeling framework belongs to the main modeling tools of contemporary mainstream macroeconomics, and the situation is unlikely to change. The use of DSGE models "in empirical work is often justified by the claim that they are the only kind of models that have the necessary resources to provide coherent theory-based stories about what happens in the economy" (Kuorikoski and Lehtinen 2018, p. 5). However, macroeconomic agent-based modeling, a framework offered by heterodox economists, may take the place of the DSGE models (cf. Colander et al. 2008).

Lawrence Christiano et al. (2011) employed the DSGE framework to investigate the effect of the so-called 'zero lower bound'[13] of the nominal interest rates on how much government spending stimulates economic development. Despite Christiano et al.'s (2011) main conclusion that "government spending multiplier can be much larger than one when the zero lower bound on the nominal interest rates binds" (p. 78) does not contain a causal-family word, it is indeed explicitly causal given that government-spending multiplier denotes the change in output (GDP) related to a change in government expenditure. For instance, the multiplier of2 depicts the situation whereby a raise of $1 in government spending changes the output by $2. Christiano et al.'s (2011) conclusion that the zero lower bound modulates the influence of government expenditure on output is based on an analysis of the DSGE model described in detail in a related paper[14] (Altig et al. 2011).

While Altig et al.'s (2011) model locates itself in the middle of DSGE models regarding its size, it is extensive and accounts for both supply and demand sides of the American economy, financial sector, and government. The model world is populated by final good and intermediary firms, households, financial intermediaries, and governments deciding on monetary and fiscal policy. For example, each household undertakes decisions regarding investment (lending capital and purchasing securities), consumption, and the expected level of wages. All the optimization decisions undertaken by these economic actors create a general equilibrium. Hence, the DSGE model is a micro-based macroeconomic model.

Interestingly, Altig et al. (2011, p. 226) employed the alternative approach to calibration compared to the previously discussed model of public transportation. Instead of choosing model parameters based on statistical data to put forward realistic assumptions (i.e., descriptively accurate of some market phenomena), the estimated values aimed at getting a model that is adequate to some empirical data without considering, in some cases, whether each parameter is plausible. Parameter values were estimated with a version of the limited information strategy (cf. Christiano et al. 2005) to make simulations based on the DSGE model resemble empirical response functions of monetary policy, technology, and capital shocks for ten crucial American macroeconomic variables within the previously estimated VAR model (cf. Chapter 3).

While this approach to calibration warrants empirical adequacy to macroeconomic data chosen for calibration, it is accused of resembling storytelling (i.e., delivering unverifiable explanations without evidentiary support). As Edward Prescott and Graham Candler (2008, p. 622) admitted, if a model aims at delivering empirical implications of a theoretical model, then "the use of statistical tools to select the parameters that best fit the business cycle observations is not sound scientific practice." The problem is that it is in principle possible to obtain in this way calibrated theoretical models adequate to data despite represented mechanisms being false (cf. Kirman 2010). In such cases, the empirical adequacy to data gives the illusion that the represented mechanism is an actual one, while it is a fallacious or, at best, possible mechanism. For instance, Joseph Stiglitz (2018) criticized the tradition of DSGE modeling for employing the assumption of rationality of economic agents despite contrary empirical evidence. If such criticism is valid (i.e., if the assumption distorts the represented mechanism in a meaningful way), then estimating model parameters allows for obtaining considerable fit to data even though the model represents a fictitious mechanism.

Christiano et al.'s (2011) analysis develops Altig et al.'s (2011) DSGE model to represent the mechanism of how the lower zero bound modulates the causal relationship between government expenditure and GDP growth. The need to construct the DSGE model arises from the fact that using direct empirical evidence (e.g., a VAR model) is impossible. As Christiano et al. (2011, p. 81) admitted, "we cannot mix evidence from states in which the zero bound binds with evidence from other states because the multipliers are very different in the two states." While the authors first construct a series of simple new-Keynesian models to represent the effect of zero lower bound on government spending multiplier, this modeling exercise delivers results in line with those of the DSGE model and, given no additional evidence[15] coming from that models, I skip it from consideration and focus on the primary tool used by the economists. The mechanism represented by such a DSGE model can be intuitively summarized as follows: under the condition of the zero lower bound, "nominal interest rate does not respond to the rise in government spending" (p. 82), and therefore, funding stimulus does not have adverse effect on the level of private investment in the economy.

Christiano et al. (2011) wanted to use Altig et al.'s (2011) DSGE model to represent and simulate the state of the American economy during the 2007–2008 financial crisis. To do so, the economists amended the numerical values of some parameters so that the model could reproduce the actual state of the economy at that time. For example, Christiano et al. (2011, p. 109 et seq.) set the value of the quarterly interest rate to be 1.0049. After modifying the previous version of the DSGE model, the economists conducted sensitivity analysis in order to understand how different conditions such as different periods of government intervention and the severity of the zero bound binding influence the estimates. This analysis indicates that both factors are positively related to the value of the multiplier; i.e., the longer the period of time when a stimulus is introduced and the longer the zero bound binds the interest rate, the more efficient is government spending in stimulating the economy.

Apart from delivering the model of mechanism that makes the lower bound on interest rates modulate the relation between government spending and output, Christiano et al. (2011, p. 114 et seq.) conducted a simulation of the American economy during and after the 2007–2008 financial crisis. This simulation seems to serve two purposes. On the one hand, the economists aim at predicting main macroeconomic variables up to 2015. On the other, the comparison of the results of the simulation to macroeconomic data seems to aim at establishing that mechanism represented by the model is an actual mechanism that governs the dynamics of the American economy. As Christiano et al. (2011) admitted, the simulation of "the model generates sensible predictions for the current crisis under the assumption that the zero bound binds. In particular, the model does well at accounting for the behavior of output, consumption, investment, inflation, and short-term nominal interest rates" (p. 82). Considering the empirical adequacy of the model prediction, it can be read as a model of an actual mechanism governing macroeconomic dynamics during recessionary demand shocks and the effects of government stimulus under the zero lower bound on interest rates. This is the case considering that the assumptions specifying the preferences of agents populating the modeled world create dynamics that replicate how the actual economy behaved during the financial crisis.

Both Christiano et al.'s (2011) DSGE model and the calibrated theoretical model previously considered have in common being a model of the mechanism. They both fulfill Marchionni's (2017) definition of an economic mechanism. However, Christiano et al.'s (2011) study overlaps with other views on causality. While the question regarding the size of the multiplier is connected to assessing average effects of manipulations, the analysis of government spending under alternative macroeconomic conditions is of counterfactual nature. Nevertheless, the mechanistic view on causality dominates the study. Otherwise, Christiano et al. (2011) would stop their analysis after constructing simpler, new-Keynesian models representing the relations among aggregated macroeconomic variables. The purpose of the DSGE model is to deliver microeconomic bases for macroeconomic causal relations; i.e., a plausible mechanism that produces the macroeconomic dependencies. These dependencies emerge from microeconomic bases (i.e., different types of agents facing optimization problems). This is visible when the economists adjust their model to replicate empirical data describing how the American economy behaved during the 2007–2008 financial crisis. Apart from estimating the value of the government spending multiplier, the DSGE model delivers a mechanistic explanation of the events following a negative shock reducing trust and risk appetite. The assumed preferences and constraints faced by consumers, intermediaries, final-good producers, the financial sector, and government lead to actions and interactions that produce the aggregate-level results similar to those observed in the target (i.e., the American economy during the crisis). This resemblance between the outcome of the mechanism operating within the model and the actual mechanism of the American economy during the 2008 recession is crucial for establishing the actuality of the modeled mechanism.

The answer to the question regarding the relata of the causal claim based on DSGE models depends on the level of reality one emphasizes. If ontological priority[16] is put on the microeconomic phenomena, then the relata of the causal claims are different types of economic agents with their preferences and constraints. On the contrary, if one prioritizes the aggregate-level phenomena, then the relata of the causal claims are variables standing for the features of macroeconomic phenomena. Semantically, Christiano et al. (2011) formulate their main causal conclusion using words denoting aggregate-level phenomena. For example, 'government spending' labels aggregate monetary value of money spent by a government in a certain period (usually reported yearly) rather than a disaggregated list of all goods and services purchased by public administration. Similarly, output denotes all products and services sold in an economy. However, considering the nature of DSGE models (i.e., micro-based macroeconomic models) leads to the conclusion that the microeconomic-level relata is a more plausible interpretation of the philosophical commitments of the modelers.

It is interesting to consider the view presupposed by Christiano et al. (2011) during the comparison of the results of simulation within the DSGE model with the variables describing the development of macroeconomic situation during the 2007–2008 financial crisis. Two remarks are in order here. First, to proceed with the comparison, the economists must believe that the mechanism represented by their DSGE model is a (possibly simplified) mechanism that produces the whole American economy. Otherwise, the influences of some other mechanisms would also influence the statistical records and therefore make the results of simulation inadequate even if the model represented an actual mechanism. Second, considering that the DSGE model represents a closed economy, such a comparison requires accepting that the influence of (geographically) external factors and international trade have an insignificant effect on the economy. Two explanations of the result (the simulated data being considerably adequate with the statistical record) are plausible. On the one hand, the DSGE model can represent the actual mechanism, and external influences have been missing in the period under study. On the other hand, the fit to data can result from the process of calibration employed by economists. Deciding which of the two is correct has severe implications for the use of this evidence for policymaking, and I will elaborate on the distinction between actual and possible mechanisms in the following Section 5.3. Here, I want to review the use of case studies for mechanistic inference. While this method is not popular in mainstream economics, related disciplines make use of it.

5.2.4 Qualitative inference of mechanisms

Case-study analyses aimed at discovering the difference-making factors are grounded in the counterfactual approach to causality (cf. Chapter 4). However, the case-study method can also be employed to theorize about possible mechanisms producing a studied phenomenon. The most popular method instantiating such an approach is 'process-tracing case study.' Despite limited use in mainstream economics, this qualitative research approach is popular in related fields, management being a prime example. Process-tracing case studies aim at tracing process, or, to put it differently, theorizing about possible causal mechanisms.[17] Such use of case-study analysis is also labeled 'process induction' defined as "the inductive observation of apparent causal mechanisms and heuristic rendering of these mechanisms as potential hypotheses for future testing" (Bennett and George 1997, p. 5). Social scientists have employed this method for a few decades, even though the methodological discussion has started recently (Kittel and Kuehn 2013). Despite the recent attention, the methods of process-tracing are poorly understood by the philosophers of social science (Collier 2011). Process tracing is practiced in several social sciences (e.g., politics, international relations), but there is a growing interest in this method among economists (Bennett and Checkel 2014). Heterodox economists (e.g., institutionalists [cf. Mirowski 2005]) and the researchers interested in the management sciences show a higher interest in the qualitative methods of causal inference.

David Collier (2011, p. 823) highlighted that the method of process tracing employed to drawing causal inferences is useful for "evaluating prior explanatory hypotheses, discovering new hypotheses, and assessing these new causal claims [. . . and] gaining insight into causal mechanisms." Bennett and Checkel (2014) describe process tracing case studies as a combination of induction and deduction. On the one hand, induction is used to draw causal (mechanistic) hypotheses from a case under consideration. On the other hand, explanations are deduced from existing theories. Therefore, the use of both approaches "depends on the prior state of knowledge and theorizing about the phenomenon and the case selected for study, and on whether the case is similar to a defined population of cases or is an outlier vis-à-vis this population" (pp. 17–18). In cases when the previously existing knowledge is limited, the inductive reasoning involves a 'time-reversed' analysis: the process is analyzed step-by-step "from the outcome of interest to potential antecedent causes" (Bennett and Checkel 2014, p. 18). Otherwise, the procedure of process tracing involves deducing possible explanation(s) of an observed process from existing theories to "develop observable implications of hypothesized mechanisms" (ibid.) and studying the case to confirm one of the hypothesized mechanisms.

The process-tracing literature delivers several 'tests' of mechanistic explanations aimed at verifying the uniqueness and certainty of a mechanical explanation. Mahoney (2012, p. 578) discussed poetically labeled 'smoking-gun tests.' Similarly to the heroes of detective stories, researchers employing the smoking-gun tests should consider if evidence indicates exclusively the operation of a particular factor or, on the contrary, it can be interpreted also in favor of another explanation. 'Hoop tests' describes testing for certainty. It is possible that a specific piece of evidence supports two mechanical (*A*, *B*) explanations, but the lack of that piece of evidence indicates that neither mechanism *A* nor mechanism *B* is at work in a studied case. In other words, failing at the hoop tests indicates that the explanation is undoubtedly false. On the contrary, passing it cannot be interpreted as supporting evidence of a high level of certitude due to the presence of other possible explanations. 'Doubly decisive tests' employ evidence that is both unique for a mechanistic explanation and certain.

On the contrary, 'straw-in-the-wind tests' are not decisive; they are neither unique nor sure, and passing such tests can only be treated as additional evidence. Some researchers advise comparing the results of process tracing to those obtained by means of other methods. For instance, Bennett and Checkel (2014, p. 19) indicated that obtaining the results that contradict other evidence lowers the likelihood of the process-tracing results to be true.

In addition to the process-tracing methodology, there are other qualitative methods of inferring mechanistic explanations of phenomena. Mahoney (2000, p. 409) analyzed the methods of causal inference employed by researchers studying small samples. One such method is within-case analysis. The following two types of such research methods are distinguished. First, 'pattern matching' is based on comparing the case under consideration with existent theories with the aim of indicating the most suitable explanation of the considered case. Second, 'causal narrative' labels the straightforward strategy of producing 'stories' of how a phenomenon under consideration is produced on the grounds of background knowledge (hitherto theories) and the case under analysis.

Furthermore, Darden (2002) discussed the following three methods of inferring mechanisms. Schema Instantiation proceeds from choosing a highly abstract type of mechanism, and adjusts it to the case under consideration. Darden exemplified this method with the case of the Darwinian mechanism that involves the production of different variants and competitive selection. Modular Subassembly is a strategy that focuses on identifying entities interacting in a considered mechanism and putting forth hypotheses about the activities these entities must undertake for a mechanism to operate. Forward/Backward Chaining is a strategy of inferring mechanisms suitable when a part of a mechanism is already known. In such cases, a researcher investigates previous (Backward Chaining) or next (Forward Chaining) stages of a causal process.

5.3 Using mechanistic evidence for policy

In this section, I discuss the use of theoretical models of mechanisms for economic policymaking. First, I argue that highly abstract, purely theoretical models can only represent possible mechanisms, and the move from a possibility claim to the actuality claim requires extensive empirical work. Furthermore, this empirical research is likely to suffice for establishing a causal claim on its own. Second, I consider whether knowledge of an actual mechanism is sufficient for conducting effective interventions. I disagree, and argue that since theoretical models represent single mechanisms, then it is in principle possible that its work will be influenced by other mechanisms operating at the same time and place. Given that mechanistic evidence does not allow for predicting the effects of interventions reliably, translating a conclusion from a mechanistic into the manipulationist understanding of causality is not justified. Finally, I find the role of mechanistic evidence in planning and introducing institutional changes that promote an expected outcome but do not warrant it.

5.3.1 From possibility to actuality: mechanist's circle

The axiomatic, deductive models that are not based on any empirical input, such as Kotchen's (2006) model of a market for green goods, have been a dominant research method in economics since 1960s until 1980s. Even though its popularity has been plummeting from then on (cf. Hamermesh 2013), theoretical models still account for one-quarter of all explicitly causal studies published in top economic journals (Maziarz 2018a). The review of the philosophical discussions concerning the models of mechanisms (cf. Section 5.1.2) indicates that such models can represent either actual or possible (i.e., not necessarily existing in the target) mechanisms. According to one of the views, the distinction between how-possibly explanations (HPE) and how-actually explanations (HAE) lies in the empirical support received by the latter (Verreault-Julien 2019). As Alisa Bokulich (2014 p. 323) put it, "a how-possibly explanation doesn't need to pick out the actual (incomplete) mechanism by which the event occurred, but it does need to be consistent with known facts. It is a candidate for a how-actually explanation."

Similarly, models of mechanisms can be interpreted either as representing possible or actual mechanisms. The latter interpretation requires empirical support for the thesis that the represented mechanism operates in the target system. The case study of Kotchen's (2006) model and earlier discussion of Schelling's (1969) toy example indicate that the empirical work needed to establish the actuality of represented mechanism is missing from such studies. Then, by default, the purely theoretical models should be interpreted as models of possible mechanisms unless further empirical evidence is delivered to support the actuality claim.

In the following chapter, I consider experimental methods of causal inference and its use as evidence for policymaking. The central problem (known as the problem of extrapolation) related to the use of experimental evidence is connected to extrapolating the results from an experimental population into a target population. Some philosophers believe that extrapolating experimental results into a policy setting requires knowledge so extensive that it is sufficient to establish a causal claim on its own. The problem is known in the philosophical literature as the extrapolator's circle (Steel 2007). Donal Khosrowi (2019, p. 45), who argued that the problem also occurs in the case of experiments in economics, defined extrapolator's circle as the situation whereby extrapolation strategies "require so much knowledge about the target population that the causal effects to be extrapolated can be identified from information about the target alone." I argue that a similar circularity appears also in the case of attempts at establishing that a purely theoretical model of a possible mechanism represents an actual mechanism.

The mechanist's circle denotes the situation when the knowledge required to establish that the mechanism represented by a model is an actual one is sufficient to support the causal claim on its own. While my argument also directly applies to the use of Kotchen's (2006) model for policy, I focus further discussion on Schelling's checkerboard model. This choice makes the discussion more intuitive. Furthermore, a considerable amount of the philosophical literature on explanation discusses this model (e.g., Aydinonat 2007; Verreault-Julien 2019). Therefore, Schelling's model can serve as a prime example of how-possibly explanation. To recall the previous discussion (cf. Section 5.1.2), the checkerboard model represents the behavior of pluses and minuses on a checkerboard and shows that a moderate preference for having neighbors with the same sign leads to strong segregation after several iterations when the pluses and minuses can exchange their positions with neighbors. This model is employed to explain the observation that American cities experience strong racial segregation despite the lack of strong preferences for such segregation.

Considering that neither Schelling (1969) nor Kotchen (2006) and other economists constructing theoretical models deliver empirical justifications for the axioms of their models, the checkerboard model (and others purely theoretical models) can represent empty mechanisms. In such cases, some other mechanism such as income distribution and spatial differentiation on the housing market is capable of producing the same effect (cf. Vinkovic and Kirman 2009). In other words, the represented mechanism may not be at work in the American cities (cf. Clark 1991), and the observed racial segregation may result from other causal factors. Based on this, Philippe Verreault-Julien (2019, p. 29) concluded that:

> we may regard the checkerboard model as providing a HPE of residential segregation in the sense that it is, in fact, possible that the preferences for not living in a minority status cause segregation. Yet, we can also consider that we have no evidence that it is the HAE.

Obviously, interventions that target an empty mechanism are doomed to failure. However, a how-possibly explanation, after receiving empirical support, can become a how-actually explanation (Brandon 1990). Using purely theoretical models as evidence of causal mechanisms requires establishing that the represented mechanism actually operates in the policy setting; i.e., the model delivers HAE.

Unfortunately, the philosophical literature lacks detailed discussions of how to assert that the represented mechanism actually operates in the world (is not empty). Considering that the outcomes and the represented mechanism follow deductively from the assumptions of the model, policymakers should be primarily concerned with establishing that the crucial assumptions[18] are realistic; i.e., describe the actual situation of the policy target. However, policymakers should also gather evidence assuring that the explanandum (i.e., the effect produced by the mechanism under consideration) is present in the policy setting. Otherwise, the modeled mechanism, while under operation, could be screened off by factors external to the model world (cf. Section 5.3.2). Finally, if the considered mechanism is screened off, other factors can produce its outcome. Hence, policymakers should also exclude the possibility that other possible (conceivable) mechanisms are not the actual ones.

Let me return to the case of Schelling's checkerboard model and its use for policymaking, and assume that a policymaker wants to counteract the racial segregation in a city. If the mechanism leading from mild preferences to the segregation is indeed actual, then modifying these preferences is one of a few interventions leading to having racially diverse districts. For example, a city could fund an educational program aimed at socializing minorities or teaching cultural differences. If the agents in Schelling's model did not prefer to live in racially unified neighborhoods, then their switches in the places where they live would not produce the segregation. Therefore, such an intervention would be successful only if the mechanism represented by the checkerboard model is an actual one. Otherwise, the intervention would not counteract the segregation. For example, if segregation is caused by income distribution and differences on housing market, then changing the preferences of inhabitants concerning the race of their neighbors would be useless.

To establish that the inhabitants of the city under consideration do indeed prefer to live in racially unified neighborhoods, policymakers have a few options at hand. They can conduct a survey and directly ask (a sample of) the residents about their preferences. However, this method can lead to fallacious results if interviewees shy away from delivering honest answers. Given that admitting to being an (even mild) racist is far from being politically correct, the results can be biased. Therefore, in this case, inferring preferences from an observation of behavior may be a better option. For example, a sociological study could involve observing how children choose their playmates. Putting methodological issues of such studies aside, let me consider further steps needed to move from possible to the actual mechanism.

Another assumption allows the agents to change their location on the checkerboard. While anecdotal evidence indicates that changing housing is possible, the process is costly (e.g., provisions, direct costs of moving, taxes, etc.), and even the members of highly mobile societies such the American society move only a few times in their lifetime. Therefore, the question of whether such a limited number of changes of location is sufficient for the mechanism to work needs to be addressed. However, obtaining evidence that the crucial assumptions of the model are accurate is not sufficient, since the conditions that produce the outcome within the model closure can fail when external factors are also at play. For example, the inhabitants of a city under consideration can prefer to have racially similar neighbors and can (sufficiently frequently) change houses, but due to some social factors (e.g., familial relations), do not do that. In such a situation, even though the crucial assumptions of the model are true, the represented mechanism is empty since external factor screens of the interactions among agents from producing the outcome that would follow under experimental closure[19] (cf. Chapter 6).

However, it is also possible that the outcome of the mechanism under consideration is produced by some other mechanism. In the case of the checkerboard model, it is possible that while city inhabitants have moderate racial preferences, they do not choose the location of their houses on this basis. Instead, they are constrained by their wealth and income level, and have strong preferences to have neighbors with a similar level of income. In such a situation, the mechanism represented by the checkerboard model is also empty. The intervention aiming at modifying the racial preferences would not result in a change in spatial differentiation because the other mechanisms would produce it. Then, the policymaker needs to make sure that the outcome is produced indeed by the mechanism under consideration.

While there are many possible approaches to ascertaining that the outcome follows from the assumptions describing the mechanism and not some other factors, estimating a cross-sectional econometric study seems to be the simplest way. Having the (mechanistic) theory describing the relation between racial preferences and spatial segregation of city inhabitants, a policymaker can construct a theory-driven model resembling case studies described in Chapter 2. The model could either utilize intra-city level data (e.g., measuring spatial segregation in different districts of one city) or use an extra-city dataset. The set of exogenous variables contains a measure of racial preferences concerning neighbors and other possible determinants (e.g., level of income, willingness to change houses, etc.). A measure of spatial segregation would be the endogenous variable in such a model. As I have argued, such a theory-driven econometric model can be interpreted as evidence for causality[20] understood in line with a version of the regularity view.

Given that the theory-driven econometric model is required to establish that the mechanism represented by the checkerboard model is an actual mechanism and establishing that the actuality is necessary to use mechanistic evidence for policymaking, then the mechanistic evidence is unnecessary to establish the causal claim. However, the presence of the mechanist's circle depends on the type of intervention under consideration. As I argue in the following Subsection 5.3.3, mechanistic evidence allows for conducting institutional reforms for which the regularity-view evidence would not be sufficient.

Furthermore, the mechanist's circle could be formulated more radically. Philosophers broadly accept the view that models are studied because their targets are inaccessible for scientists. This claim can have different formulations. For instance, Jessica Bolker (2009) studied physical models used in biology and differentiated between exemplary (being a small sample of a group of organisms) and surrogate models. Biologists use the latter type of models due to inaccessibility of their targets. For example, they experiment on rodents because conducting experiments on human animals is considered unethical. In economics, the inaccessibility is usually of epistemic or economic nature.

Under this view, theoretical models are constructed and studied by economists because a direct analysis of their targets is impossible: "the target systems are inaccessible in full since they are too small, too large, too far away in space or time or too complex" (Mäki 2005, p. 304). The surrogate systems (i.e., theoretical models in this case) are used because experimenting on the whole economy is either too costly or simply impossible (imagine randomizing countries and causing hyperinflation in the treatment group to study the effects of printing money). As Uskali Mäki (2005, p. 304) put it, "[t]he epistemic point of this activity is that the properties of such substitute or surrogate systems are directly examined in order to indirectly acquire information about the properties of the system they represent." If the view that theoretical models are used due to the epistemic inaccessibility of their targets, then the move from the interpretation of a model of possible mechanism to the claim about actuality is in principle impossible.

5.3.2 Actual mechanisms and external influences

In this section, I argue that mechanistic knowledge cannot be translated into the manipulationist notion due to the external influences that are likely to make the effects of interventions differ from what is expected. Despite Federica Russo and Jon Williamson's (2007, p. 162) opinion that "[i]t is uncontroversial that mechanistic evidence on its own cannot warrant a causal claim, as it may be the case that the purported cause, although prior to the effect and mechanistically connected to it, actually makes little or no difference to it," theoretical models that, as I have argued above, represent single mechanisms, are often appraised and considering as delivering justification for policymaking.

The presence of the mechanist's circle undermines using models of possible mechanisms as evidence for some types of policymaking. However, other types of mechanistic studies establish the actuality of mechanisms using calibration. In such cases, the mechanist's circle does not apply. However, if a policymaker wants to use mechanistic evidence with the actuality claims to intervene in a way that requires translating the meaning of causal claim from the mechanistic into the manipulationist notion (i.e., by changing the relata), the problem of external influences occurs: the factors that are excluded from the model may distort the effects of policymaking.

Mechanistic knowledge gives the illusion that one can predict what the effects are of an intervention modifying the conditions of agents' choices. This indeed is the case if the represented mechanism is isolated from external influences. Think of the case of Robinson Crusoe considered at first-year economics courses. The Robinson Crusoe model represents consumer and producer choices in the economy with only one consumer, one producer, and two goods (cf. Varian 2014, p. 629 et seq.). If a policymaker wanted to intervene on such a desert island (e.g., by delivering the consumer good), then the effects of this intervention (higher consumption and lower production) could be deduced from the model. However, these conclusions hold only within the model world and in a world that resembles (to a sufficiently high degree) the model concerning the isolation of the mechanism from external influences. If a policymaker used the model of the Robinson Crusoe economy as evidence for intervention in the American economy, the deduced effects would likely not follow due to other (extra-model) factors. For instance, the existence of substitute goods would change Robinson's behavior.

Previously, I have studied three examples of theoretical models. Kotchen's (2006) purely theoretical model of green good markets isolates the influence of possible determinants of the demand for such goods from many factors. The model represents the mechanism of demand for green goods. Macroeconomic conditions are the most straightforward example of factors excluded from the model. Such isolations serve the purpose of simplicity; i.e., making the representation of mechanism epistemically possible.

Similar to the model of Robinson Crusoe economy considered previously, an intervention on one of the factors considered by Kotchen (2006) may result in very different (or even opposite) outcomes due to external influences such as macroeconomic crisis or technological shock. Similarly, Parry and Small's (2009) calibrated theoretical model represents the mechanism of how subsidies promote the use of public transport. However, the mechanism can be influenced in many different ways by external factors (e.g., by an atomic bomb or a macroeconomic slowdown that acts on people's need to commute to work).

The situation may be different in the case of Christiano et al.'s (2011) DSGE model. This class of model is sufficiently advanced to represent an (idealized) mechanism of the whole economy. However, the DSGE models, despite their complexity, also exclude some factors from analysis. For example, Christiano et al. (2011) assumed the American economy to be a closed system (i.e., they isolated away international trade). Although the comparison of simulated path with macroeconomic time series from the crisis suggests that the influences of external factors (such as foreign trade) are omittable, this assumption is obviously false, and even if the changes in the conditions of foreign trade did not influence the macroeconomic variables in the period under comparison, they might do so in the future.

In summary, all models of mechanisms represent mechanisms isolated from some factors. Given that theoretical models represent mechanisms isolated from external influences and policymakers undertake their actions in the real world and not the model world, then it is in principle possible that some external influences will distort the effects of policymaking. In order to predict the effects of interventions conducted in the real world and not in the model world, policymakers would need to have knowledge of all mechanisms operating in the economy and, additionally, knowledge of the interactions among these mechanisms. While the recent work on multiple models in economics (Rodrik 2015; Aydinonat 2018) gives hope that such knowledge may be obtained in the future, the current state of economics does not allow for the translation from the mechanistic into the manipulationist notion of causality.

5.3.3 The advantage of mechanistic evidence

The previous arguments shed the light of skepticism on the role of theoretical models in economic research and, particularly, the use of such evidence for policymaking. First, theoretical models usually establish only possibility claims and require extensive empirical work to confirm the actuality of represented mechanisms. Usually, this work on its own suffices for producing a causal claim. Furthermore, considering that many mechanisms operate in the world of the economy at the same time while models represent only single mechanisms, predicting the outcome of interventions using only one model of mechanism is impossible. However, mechanistic evidence can play an important role in policymaking because it delivers knowledge not present in causal claims presupposing other views on causality. Mechanistic knowledge (even the knowledge of possible mechanisms) allows for deducing how changes in agents' preferences and constraints influence their choices. Therefore, mechanistic evidence can be used to modify these constraints and preferences with a view to promote (rather than obtain with some degree of certitude) policymakers' goals.

Such a change is neither intervention in the strict sense (i.e., a change in a relatum of a causal claim), nor the use of (possibly misleading) causal structure to act on the basis of predictions. Institutional[21] reforms not necessarily lead to expected outcomes (due to external influences) but, appropriately introduced, they promote a policy target and make it more likely to happen. As I argue ahead, institutional reforms can be based on models of possible mechanisms if they create the mechanism in the target. Let me consider an imaginary institutional reform based on Kotchen's (2006) model of a possible mechanism for the market of green goods. One of the conclusions of the model is that consumers should (under some further conditions) prefer to choose a combination of green goods and either purely private or purely public goods if the green good is more efficient in delivering both private and public needs. To restate the summary of Kotchen's (2006) paper from the previous section, green goods are such products that fulfill two types of needs. Two other products in Kotchen's model world are purely private goods that fulfill the 'private' need and purely public goods that fulfill the public need. For example, you can think of private need as a need to commute and public need as a need to reduce the carbon dioxide emission. In such a case, electric cars can instantiate Kotchen's (2006) green good.

I have argued that the 'technological assumption' stating that the green good is more efficient in fulfilling both needs than the purely private and purely public good is the crucial assumption. If the opposite were assumed, then the conclusions of the model would be reversed: rational consumers, in that case, would purchase a combination of purely private and purely public goods, and the demand for the green good would equal zero. Let me assume that the policymaker wants to promote the purchase of the green good for some reason.[22] To do so, they can *make* the technological assumption a true description of their policy setting. Introducing an ecological tax for the purely private good, promoting technological advancement of the green good, or subsidizing the production of the green good are three exemplary options of introducing such a policy.

However, the institutional reform relying on making the technological assumption true does not guarantee that the result will follow. The intervention makes the result more likely, but does not warrant success. This is the case for the reasons I have considered in previous sections. First, there is the question of whether the other parts of the mechanism which are present stay open when a policymaker uses a model of a possible mechanism as evidence for the institutional change. For example, institutional reform can fail at being successful because consumers have only preferences that are fulfilled by purely private good (i.e., they are selfish) and will choose the purely private good (or its real-world counterparts) regardless of price. Second, even if an institutional reform makes a possible mechanism operate in the target (make it actual), other causal mechanisms operating at the same time can interact with or even screen off the mechanism from producing the effect that could be expected on the basis of the model.

Let me exemplify these conclusions with the case of institutional policy introducing a tax exemption[23] promoting purchases of electric cars. As I have argued previously, the obstacles that can prevent the reform from being efficient can either result from the other parts of the mechanism being absent or other mechanisms influencing the outcome. First, it is easy to imagine that despite the institutional reform targeting the technological assumption of Kotchen's (2006) model, the represented mechanism stays possible. For example, it may be the case that a consumer prefers to purchase traditional cars and support planting trees, even though it is more efficient to buy a subsidized electric car. Second, other factors (excluded from Kotchen's [2006] model), such as the use of other modes of transportation, can modify consumers' preferences so that they choose to refrain from buying any car. As the example and the stated considerations show, institutional reforms can influence economic agents' constraints and, indirectly, preferences to promote an outcome, but be unable to warrant that the result will follow.

5.4 Mechanisms as evidence for institutional reforms

The mechanistic view on causality states that what distinguishes accidental and causal relationships is that processes or mechanisms produce the latter. The definitions present in the philosophical literature have in common defining mechanisms as parts and direct interactions between the parts of a system that produce higher-level dependences. Economists delivering causal claims on the basis of theoretical, deductive models accept a version of the mechanistic view on causality that identifies the 'parts of a system' with economic agents. Consumers, companies, governments, and financial institutions are exemplary agents. These agents face constraints and undertake optimization decisions. Their actions and the interactions among agents produce causal dependencies among aggregate (either microeconomic or macroeconomic) variables.

Economists use different types of theoretical, deductive models to represent such mechanisms. These axiomatic models can be divided into the following two main categories. First, purely theoretical models are highly abstract and are used to represent idealized and isolated possible mechanisms. Kotchen's (2006) theoretical study of markets for green goods is a representative example of such modeling exercises. Second, calibrated theoretical models belong also to the group of deductive models, but they also use empirical input: the values of parameters are adjusted to describe markets that are targets of these models. Parry and Small's (2009) study of the influence of subsidies on the use of public transport instantiates this type of models. The class of DSGE models is a special case of calibrated models. They deserve separate treatment due to their widespread use for macroeconomic policymaking. Christiano et al.'s (2011) model of the influence of zero lower bound on the effectiveness of fiscal policy is an example of this type of macroeconomic model. Furthermore, Christiano et al.'s (2011) model serves as an example of how the move from a possibility claim to the actuality claim can be made. The DSGE modelers calibrated the assumptions of their model to represent (in a simplified way) the American economy and run a simulation showing that the model is able to reproduce the behavior of the economy during the 2007–2008 financial crisis.

Two problems occur when the results of theoretical modeling are used as evidence for policymaking. Theoretical models usually represent possible mechanisms, and it is policymakers' responsibility to gather empirical evidence sufficient for the move from the possibility claim into the actuality claim. Unfortunately, as I have argued, the empirical evidence required for this move is extensive enough that it is sufficient to establish a causal claim on its own. Therefore, similarly to the extrapolator's circle discussed in the following chapter, the mechanist's circle can undermine the use of mechanistic evidence for policymaking. The status of theoretical models of single mechanisms resembles the problem with extrapolating the results of experiments (Chapter 6). Both theoretical models and experiments exclude some factors, and taking causal the claims true in a model world as evidence for interventions in the actual world may lead to erroneous results.

Furthermore, I have argued that the effects of interventions relying on changing the relata of (mechanistic) causal claims may differ from expected outcomes. This is the case because (1) theoretical models usually represent only one (or a few) mechanisms, while (2) there are many mechanisms that operate in the world of economy, and therefore (3) the work of one (modeled) mechanism used as evidence for intervention may be disrupted by external influences. The presence of many mechanisms operating at the same time in the world makes the translation of causal claims from the mechanistic notion into the manipulationist notion unwarranted. However, policymakers can employ mechanistic evidence (even describing possible mechanisms) to conduct institutional reforms. This type of policymaking relies on modifying agents' constraints or preferences to make crucial assumptions of a model true about a target. Institutional reforms can make a mechanism more likely to operate within the target system, but they cannot warrant the outcome that would follow within the model world to be produced by that mechanism in the actual world.

Notes

1 Darden's definition of mechanism is formulated as interactions of entities that, due to having certain properties, behave (are active) in given ways.
2 The original, German-language title is *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie.*
3 Schelling's (1969) purpose was not to explain any phenomena whatsoever, but to construct a simple and mathematically tractable model for educational purposes.
4 The most simplified version of the checkerboard model is in fact constructed in one dimension. Hence, contrary to the name, there is a line instead of the 'checkerboard.'
5 Earlier DSGE models employed 'representative agent' standing for different types of entities of the modeled mechanism.
6 Impure public goods are such products and services that have jointly the features of private goods and public goods. The study exemplifies impure public goods with 'green products'; i.e., the more expensive versions of goods that are produced in a more ecological way. For instance, buying 'shade coffee beans' (i.e., the beans from the plantations located under rainforest instead of replacing it) delivers both espresso (private good) and biodiversity (public good).
7 How-possibly explanations (i.e., delivering a model of a mechanism that could produce the observed outcome without establishing that this mechanism was indeed at work in a given situation) are in stark contrast with how-actually explanations that not only deliver a possible mechanism but also show that the modeled mechanism has actually produced the explained outcome. One of the views on the sources of difference between models delivering how-actually and how-possibly explanations is that it lies in the empirical support lacked by the latter (e.g., Hempel 1965).
8 Even though philosophers and economists discussed the role of assumptions and their realism, among other topics, the literature remains silent on the question of how modelers choose the assumptions. For instance, do they choose axioms that allow for obtaining an intuitive result, or consider the realism of crucial assumptions on the basis of empirical research?
9 Today, the mainstream economics discusses preferences and refrains from measuring utility but the view that the consumption of different goods delivers varying amounts of 'utils,' that has been put forward by Jeremy Bentham (2007 [1789]), had once been popular in economics.
10 I need to highlight that the purpose of this discussion is not to indicate the scope of the model or establish the actuality of the mechanism, but only to deliver examples for the argument that the technological assumption, which is crucial for the results obtained by Kotchen (2006), can possibly be in disagreement with empirical findings.
11 Considering the complication of the model and the purpose of the book, I refrain from discussing all assumptions in detail. An interested reader can easily find such description in the original paper (Parry and Small 2009, pp. 706–711).
12 Numéraire denotes a (theoretical or actual) good that is used in an economic model to denote the values of other goods (e.g., Brekke 1997). In this case, the representative agent resigns from the consumption of numéraire (interpreted as all other goods) in exchange for using more expensive mean or transportation or commuting more often.
13 In response to the 2007–2008 financial crisis and subsequent recession, several central banks cut interest rates so much that it reached the level of 0%. While we now know that the central bank interest rates can be set at a level below 0 (what implies that holding bank deposits is taxed to promote investment and spending), the zero bound has been

considered to be the lower limit. Furthermore, some economists still argue that such a strong incentive against saving has adverse effects in the long term (e.g., Palley 2016).

14 In response to the dominance of 'publish or perish' culture in the environment of academic economists, dividing the results of one's work among several independent outputs is a common practice aimed at boosting publication record – and hence, chances for promotion (Moosa 2018).

15 Considering that these new-Keynesian models are models of functional dependencies among aggregate variables, the DSGE model that represents mechanism producing the dependencies delivers stronger evidence for the causal claim.

16 Assuming a simplifying division of economic reality into only two levels (i.e., micro and macro) allows for formulating two different answers to the question regarding the relata of causal claims. By 'ontological priority,' I mean putting emphasis or focusing on one of these two.

17 It should be highlighted that not every use of the process-tracing method is a case of the causal inference aimed at discovering causal mechanism. This approach to single-case studies can also be descriptive in nature. Therefore, it is up to an investigator whether conclusions are purely descriptive or causal in nature.

18 Obviously, assumptions in a theoretical model play different roles. Some assumptions describe the mechanism represented by a model. Others isolate it from external influences. Still others idealize (e.g., simplify) the modeled mechanism. While some assumptions not necessarily have to be accurate for a represented mechanism to be actual, others play a crucial role. For example, the assumption of the more simplified version of the checkerboard model locating pluses and minuses in a one-dimensional space is obviously false if it is taken as a description of the geography of American cities. However, despite idealizing the mechanism by simplification, it does not necessarily distort it in a meaningful way, since a similar mechanism could also operate in a three-dimensional space. Other assumptions, such as the axiom stating that agents have mild preferences towards racial segregation, are crucial in the sense that if they are not accurate concerning the policy setting, then the mechanism is an empty one.

19 Experimental designs, especially in the natural sciences, isolate one or a few factors from external influences; i.e., close the studied phenomenon (hence, experimental closure). For example, to study experimentally the gravitational force, one needs to artificially construct a vacuum. Such a research design isolates gravitation from the friction produced by air particles.

20 Here, I assume that the model is sufficiently accurate to data and the parameter showing the partial correlation between the level of racial views in a district and spatial segregation is positive and significant.

21 In economics, 'institution' refers not only to what is meant by this word in the everyday language (e.g., companies, government agencies, etc.) but to 'rules and practices of the game' of economic exchange (cf. Duina 2013, p. 2). Hence, institutional reforms is such policymaking that modifies these rules).

22 The consideration of externalities of consuming different types of goods can deliver the reason for such a policy. For example, reducing carbon dioxide emissions can be obtained either by switching to an electric car or using a traditional one and planting trees. While both options can be equally efficient in reducing greenhouse gases, the former option may be preferable from the perspective of the policymaker with a view to reduce other emissions (such as particulate matter).

23 For a review of fiscal incentives introduced to promote purchases of electric vehicles in the European Union, see Góomez Vilchez and Thiel (2019).

References

Alexandrova, A., & Northcott, R. (2013). It's just a feeling: Why economic models do not explain. *Journal of Economic Methodology*, *20*(3), 262–267. DOI: 10.1080/1350178X.2013.828873

Altig, D., Christiano, L. J., Eichenbaum, M., & Linde, J. (2011). Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic Dynamics*, *14*(2), 225–247. DOI: 10.1016/j.red.2010.01.001

Aydinonat, N. E. (2007). Models, conjectures and exploration: An analysis of Schelling's checkerboard model of residential segregation. *Journal of Economic Methodology*, *14*(4), 429–454. DOI: 10.1080/13501780701718680

Aydinonat, N. E. (2018). The diversity of models as a means to better explanations in economics. *Journal of Economic Methodology*, *25*(3), 237–251. DOI: 10.1080/1350178X.2018.1488478

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *36*(2), 421–441. DOI: 10.1016/j.shpsc.2005.03.010

Bechtel, W., & Richardson, R. (2010). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research.* Cambridge: The MIT Press. DOI: 10.7551/mitpress/8328.001.0001

Beck, N. (2006). Is causal-process observation an oxymoron? *Political Analysis*, *14*(3), 347–352. DOI: 10.1093/pan/mpj015

Bennett, A., & Checkel, J. (2014). *Process Tracing.* Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139858472

Bennett, A., & George, A. L. (1997). *Process Tracing in Case Study Research.* Washington, DC: MacArthur Foundation.

Bentham, J. (2007 [1789]). *An Introduction to the Principles of Morals and Legislation.* North Chelmsford: Courier Corporation.

Bhaskar, R. (2008). *A Realist Theory of Social Science.* Abingdon: Routledge.

Bokulich, A. (2014). How the tiger bush got its stripes: 'How possibly' vs. 'how actually' model explanations. *The Monist*, *97*(3), 321–338. DOI: 10.5840/monist201497321

Bolker, J. A. (2009). Exemplary and surrogate models: Two modes of representation in biology. *Perspectives in Biology and Medicine*, *52*(4), 485–499. DOI: 10.1353/pbm.0.0125

Brandon, R. (1990). *Adaptation and Environment.* Princeton: Princeton University Press.

Broadbent, A. (2011). Inferring causation in epidemiology: Mechanisms, black boxes, and contrasts. In: Illari, Ph., Russo, F., & Williamson, J. (eds.) *Causality in the Sciences* (pp. 45–69). Oxford: Oxford University Press.

Brekke, K. A. (1997). The numéraire matters in cost-benefit analysis. *Journal of Public Economics*, *64*(1), 117–123. DOI: S0047-2727(96)01610-6

Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties*, *2*(1), 11–20. DOI: 10.1017/S1745855207005029

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, *113*(1), 1–45. DOI: 10.1086/426038

Christiano, L., Eichenbaum, M., & Rebelo, S. (2011). When is the government spending multiplier large? *Journal of Political Economy*, *119*(1), 78–121. DOI: 10.1086/659312

Clark, W. A. (1991). Residential preferences and neighborhood racial segregation: A test of the Schelling segregation model. *Demography*, *28*(1), 1–19.

Colander, D., Howitt, P., Kirman, A., Leijonhufvud, A., & Mehrling, P. (2008). Beyond DSGE models: Toward an empirically based macroeconomics. *American Economic Review*, *98*(2), 236–40. DOI: 10.1257/aer.98.2.236

Collier, D. (2011). Understanding process tracing. *PS: Political Science & Politics*, *44*(4), 823–830. DOI: 10.1017/S1049096511001429

Craver, C., & Tabery, J. (2017). Mechanisms in science. In: Zelta, E. (ed.) *The Stanford Encyclopedia of Philosophy: 2017 Edition.* Spring. Retrieved from: https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/. Access: 13th March 2018.

Darden, L. (2002). Strategies for discovering mechanisms: Schema instantiation, modular subassembly, forward/backward chaining. *Philosophy of Science*, *69*(S3), S354–S365. DOI: 10.1086/341858

Dawkins, Ch. (2001). Calibration. In: Heckman, J. & Leamer, E. (eds.) *Handbook of Econometrics* (pp. 3653–3703). Amsterdam: North Holland.

Del Negro, M., & Schorfheide, F. (2013). DSGE model-based forecasting. In: *Handbook of Economic Forecasting* (Vol. 2, pp. 57–140). Elsevier. DOI: 10.1016/B978-0-444-53683-9.00002-5

Dowe, P. (2001). A counterfactual theory of prevention and 'causation' by omission. *Australasian Journal of Philosophy, 79*(2), 216–226.

Dowe, P. (2012). Causal process theories. In: Beebee, H. et al. (eds.) *The Oxford Handbook of Causation* (pp. 213–234). Oxford: Oxford University Press.

Duina, F. (2013). *Institutions and the Economy*. Cambridge: John Wiley and Sons.

Elster, J. (1983). *Explaining Technical Change: A Case Study in the Philosophy of Science*. Cambridge: Cambridge University Press.

Fernández-Villaverde, J. (2010). The econometrics of DSGE models. *SERIEs*, *1*(1–2), 3–49.

Friedman, M. (1953). *Essays in Positive Economics*. Chicago: University of Chicago Press.

Frigg, R., & Hartmann, S. (2006). Scientific models. In: Sarkar, S. & Pfeifer, J. (eds.) *The Philosophy of Science: An Encyclopedia*. London: Routledge.

Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, *98*(2), 341–354. DOI: 10.1017/S0003055404001182

Gerring, J. (2010). Causal mechanisms: Yes, but…. Comparative political studies, *43*(11), 1499–1526. DOI: 10.1177/0010414010376911

Glaister, S. (1974). Generalised consumer surplus and public transport pricing. *Economic Journal*, *84*(336), 849–867.

Glennan, S. S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, *44*(1), 49–71.

Glennan, S. S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, *69*(S3), S342–S353.

Glennan, S. S. (2010). Ephemeral mechanisms and historical explanation. *Erkenntnis*, *72*(2), 251–266. DOI: 10.1086/341857

Glennan, S. S. (2012). Mechanisms. In Beebee, H., Hitchcock, C. & Menzies, P. (eds.) *The Oxford Handbook of Causation* (pp. 315–325). Oxford: Oxford University Press.

Gómez Vilchez, J. J., & Thiel, C. (2019). The effect of reducing electric car purchase incentives in the European Union. *World Electric Vehicle Journal*, *10*(4), 64. DOI: 10.3390/wevj10040064

Greenberg, R., Bichier, P., & Sterling, J. (1997). Bird populations in rustic and planted shade coffee plantations of Eastern Chiapas, Mexico. *Biotropica*, *29*(4), 501–514. DOI: 10.1111/j.1744-7429.1997.tb00044.x

Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis*, *70*(1), 81–99.

Grüne-Yanoff, T. (2016). Why behavioural policy needs mechanistic evidence. *Economics & Philosophy*, *32*(3), 463–483. DOI: 10.1017/S0266267115000425

Grüne-Yanoff, T., & Marchionni, C. (2018). Modeling model selection in model pluralism. *Journal of Economic Methodology*, *25*(3), 265–275. DOI: 10.1080/1350178X.2018.1488572

Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, *51*(1), 162–172. DOI: 10.1257/jel.51.1.162

Hardt, Ł. (2014). Modele mechanizmów i ich rola w wyjaśnianiu w ekonomii. *Ekonomia. Rynek, Gospodarka, Społeczeństwo*, *37*, 79–103.

Hardt, L. (2016). The recent critique of theoretical economics: A methodologically informed investigation. *Journal of Economic Issues*, *50*(1), 269–287. DOI: 10.1080/00213624.2016.1148984

Hardt, Ł (2017). *Economics without Laws*. Cham: Palgrave Macmillan.

Hartwig, M. (2015). *Dictionary of Critical Realism*. London: Routledge.

Hausman, D. (1992). *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.

Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century retrospective. *The Quarterly Journal of Economics*, *115*(1), 45–97. DOI: 10.1162/003355300554674

Hempel, C. G. (1965). Aspects of scientific explanation. In: *Aspects of Scientific Explanation: And Other Essays in the Philosophy of Science* (pp. 331–497). New York: Free Press.

Henderson, V. J. (1977). *Economic Theory and the Cities*. New York: Academic Press.

Hodgson, G. M. (2004). Darwinism, causality and the social sciences. *Journal of Economic Methodology*, *11*(2), 175–194. DOI: 10.1080/13501780410001694118

Iacoviello, M. (2005). House prices, borrowing constraints, and monetary policy in the business cycle. *American Economic Review*, *95*(3), 739–764. DOI: 10.1257/0002828054201477

Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*(1), 119–135. DOI: 10.1007/s13194-011-0038-2

Jansson, J. O. (1979). Marginal cost pricing of scheduled transport services: A development and generalisation of Turvey and Mohring's theory of optimal bus fares. *Journal of Transport Economics and Policy*, *13*(3), 268–294.

Khosrowi, D. (2019). Extrapolation of causal effects: Hopes, assumptions, and the extrapolator's circle. *Journal of Economic Methodology*, *26*(1), 45–58. DOI: 10.1080/1350178X.2018.1561078

Kirman, A. (2010). The economic crisis is a crisis for economic theory. *CESifo Economic Studies*, *56*(4), 498–535. DOI: 10.1093/cesifo/ifq017

Kittel, B., & Kuehn, D. (2013). Introduction: Reassessing the methodology of process tracing. *European Political Science*, *12*(1), 1–9. DOI: 10.1057/eps.2012.4

Kotchen, M. J. (2006). Green markets and private provision of public goods. *Journal of Political Economy*, *114*(4), 816–834. DOI: 10.1086/506337

Krajewski, W. (1997). Energetic, informational, and triggering causes. *Erkenntnis*, *47*(2), 193–202. DOI: 10.1023/A:1005313904241

Krugman, P. (1993). How I work. *The American Economist*, *37*(2), 25–31. DOI: 10.1177/056943459303700204

Krugman, P. (2009). How did economists get it so wrong? *New York Times*. Retrieved from: www.nytimes.com/2009/09/06/magazine/06Economic-t.html?_r=1&pagewanted=print. Access: 18th March 2019.

Kuorikoski, J., & Lehtinen, A. (2018). Model selection in macroeconomics: DSGE and ad hocness. *Journal of Economic Methodology*, *25*(3), 252–264. DOI: 10.1080/1350178X.2018.1488563

Kuorikoski, J., Lehtinen, A., & Marchionni, C. (2010). Economic modelling as robustness analysis. *The British Journal for the Philosophy of Science*, *61*(3), 541–567. DOI: 10.1093/bjps/axp049

Latour, B., & Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton: Princeton University Press.

Lawson, T. (1997). *Economics and Reality*. London: Routledge. DOI: 10.4324/9780203195390

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25. DOI: 10.1086/392759

Mahoney, J. (2000). Strategies of causal inference in small-N analysis. *Sociological Methods & Research*, *28*(4), 387–424. DOI: 10.1177/0049124100028004001

Mahoney, J. (2012). The logic of process tracing tests in the social sciences. *Sociological Methods & Research*, *41*(4), 570–597. DOI: 10.1177/0049124112437709

Mäki, U. (2005). Models are experiments, experiments are models. *Journal of Economic Methodology*, *12*(2), 303–315. DOI: 10.1080/13501780500086255

Mäki, U. (2006 [1992]). On the method of isolation in economics. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, *26*, 19–54.

Mäki, U. (2008). Realism. In: Hausman, D. (ed.) *The Philosophy of Economics: An Anthology* (pp. 431–439). Cambridge: Cambridge University Press.

Mäki, U. (2009a). Missing the world: Models as isolations and credible surrogate systems. *Erkenntnis*, *70*(1), 29–43. DOI: 10.1007/s10670-008-9135-9

Mäki, U. (2009b). Realistic realism about unrealistic models. In: Kincaid, H. & Ross, D. (eds.) *Oxford Handbook of the Philosophy of Economics* (pp. 68–98). Oxford: Oxford University Press.

Mäki, U. (2009c). Models and truth: The functional decomposition approach. In: *European Philosophy of Science*. Cham: Springer.

Mäki, U. (2011a). Models and the locus of their truth. *Synthese*, *180*(1), 47–63. DOI: 10.1007/s11229-009-9566-0

Mäki, U. (2011b). Scientific realism as a challenge to economics (and vice versa). *Journal of Economic Methodology*, *18*(1), 1–12. DOI: 10.1080/1350178X.2011.553372

Mäki, U. (2013). On a paradox of truth, or how not to obscure the issue of whether explanatory models can be true. *Journal of Economic Methodology*, *20*(3), 268–279. DOI: 10.1080/1350178X.2013.828869

Mäki, U., & Caldwell, B. J. (1992). The market as an isolated causal process: A metaphysical ground for realism. In: *Austrian Economics: Tensions and New Directions* (pp. 35–65). Dordrecht: Springer.

Marchionni, C. (2017). Mechanisms in economics. In: Glennan, S. & Illari, Ph. (eds.) *The Routledge Handbook of Mechanisms and Mechanical Philosophy* (pp. 423–434). London: Routledge.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, *33*(2), 3–11. DOI: 10.3102/0013189X033002003

Maziarz, M. (2018a). Causal inferences in contemporary economics. *Mendeley Data*. Retrieved from: http://doi.org/10.17632/v7dhjnd8xg.2. Access: 22nd March 2019.

Maziarz, M. (2018b). *Disentangling the Philosophy of Economics*. Warsaw: The Publishing House of the Institute for Consumption, Market, and Business Cycle Research.

Mirowski, P. (2005). The philosophical bases of institutionalist economics. In: *Economics and Hermeneutics* (pp. 85–121). Abingdon: Routledge.

Mishkin, F. S. (1996). *The Channels of Monetary Transmission: Lessons for Monetary Policy* (No. w5464). Cambridge: National Bureau of Economic Research.

Moneta, A., & Russo, F. (2014). Causal models and evidential pluralism in econometrics. *Journal of Economic Methodology*, *21*(1), 54–76. DOI: 10.1080/1350178X.2014.886473

Moosa, I. A. (2018). *Publish or Perish: Perceived Benefits Versus Unintended Consequences*. Edward Elgar Publishing. DOI: 10.17159/1727-3781/2018/v21i0a5441

Morgan, M. S., & Knuuttila, T. (2012). Models and modelling in economics. In: *The Oxford Handbook of the Philosophy of Economics* (pp. 49–87). Oxford: Oxford University Press.

Palley, T. I. (2016). Why Negative Interest Rate Policy (NIRP) is ineffective and dangerous. *Real-World Economics Review*, *76*, 5–15.

Papineau, D. (1996). *The Philosophy of Science*. Oxford: Oxford University Press.

Parry, I. W., & Small, K. A. (2009). Should urban transit subsidies be reduced? *American Economic Review*, *99*(3), 700–724. DOI: 10.1257/aer.99.3.700

Prescott, E. C., & Candler, G. V. (2008). Calibration. In: *The New Palgrave Dictionary of Economics* (Vol. 1–8, pp. 622–625). London: Palgrave Macmillan.

Racz, A. A., Muntean, I., & Stan, S. D. (2015). A look into electric/hybrid cars from an ecological perspective. *Procedia Technology*, *19*, 438–443. DOI: 10.1016/j.protcy.2015.02.062

Railton, P. (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science*, *45*(2), 206–226. DOI:10.1086/288797

Reiss, J. (2007). Do we need mechanisms in the social sciences? *Philosophy of the Social Sciences*, *37*(2), 163–184. DOI: 10.1177/0048393107299686

Reiss, J. (2012). The explanation paradox. *Journal of Economic Methodology*, *19*(1), 43–62. DOI: 10.1080/1350178X.2012.661069

Reiss, J. (2013). Contextualising causation part I. *Philosophy Compass*, *8*(11), 1066–1075. DOI: 10.1111/phc3.12074

Rodrik, D. (2015). *Economics Rules: The Rights and Wrongs of the Dismal Science*. WW Norton & Company. DOI: 10.17323/1726-3247-2015-4-39-59

Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, *21*(2), 157–170. DOI: 10.1080/02698590701498084

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Salmon, W. C. (1994). Causality without counterfactuals. *Philosophy of Science*, *61*(2), 297–312. DOI: 10.1086/289801

Sbordone, A. M., Tambalotti, A., Rao, K., & Walsh, K. J. (2010, October). Policy analysis using DSGE models: An introduction. *Economic Policy Review*, 23–43. DOI: 10.2139/ssrn.1692896

Schelling, Th. (1969). Models of segregation. *American Economic Review: Papers and Proceedings*, *59*(2), 488–493.

Schelling, Th. (1978). *Micromotives and Macrobehavior*. New York: Norton.

Smets, F., & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, *1*(5), 1123–1175. DOI: 10.1162/154247603770383415

Smith, A. (1950 [1776]). *An Inquiry Into the Nature and Causes of the Wealth of Nations*. Methuen. Retrieved from: website.aub.lb/fas/Documents/Adam-SMITH.doc

Steel, D. (2004). Social mechanisms and causal inference. *Philosophy of the Social Sciences*, *34*, 55–78. DOI: 10.1177/0048393103260775

Steel, D. (2007). *Across the Boundaries: Extrapolation in Biology and Social Science*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195331448.001.000

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, *34*(1–2), 70–106. DOI: 10.1093/oxrep/grx057

Sugden, R. (2000). Credible worlds: The status of theoretical models in economics. *Journal of Economic Methodology*, 7(1), 1–31. DOI: 10.1080/135017800362220

Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis*, *70*(1), 3–27. DOI: 10.1007/s10670-008-9134-x

Thünen, von J., Tribe, K., & Suntum, U. (2009). *The Isolated State in Relation to Agriculture and Political Economy*. Hampshire: Palgrave Macmillan.

Ullmann-Margalit, E. (1978). Invisible-hand explanations. *Synthese*, *39*(2), 263–291. DOI: 10.1007/BF00485077

Valdivia, D. (2015). Handbook on DSGE models: Some useful tips in modeling a DSGE models. Retrieved from: https://dx.doi.org/10.2139/ssrn.2550303. Access: 18th April 2019.

Varian, H. R. (2014). *Intermediate Microeconomics with Calculus: A Modern Approach*. WW Norton & Company.

Verreault-Julien, P. (2017). Non-causal understanding with economic models: The case of general equilibrium. *Journal of Economic Methodology*, *24*(3), 297–317. DOI: 10.1080/1350178X.2017.1335424

Verreault-Julien, P. (2019). How could models possibly provide how-possibly explanations? *Studies in History and Philosophy of Science Part A*, *73*, 22–33. DOI: 10.1016/j.shpsa.2018.06.008

Vinkovic, D., & Kirman, A. (2009). *Schelling's Model with Income Preferences and a Housing Market*. Marseille: GREQUAM.

Williamson, J. (2011a). Mechanistic theories of causality part I. *Philosophy Compass*, *6*(6), 421–432. DOI: 10.1111/j.1747-9991.2011.00400.x

Williamson, J. (2011b). Mechanistic theories of causality part II. *Philosophy Compass*, *6*(6), 433–444. DOI: 10.1111/j.1747-9991.2011.00401.x

Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic Monthly*, *249*(3), 29–38.

Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, *69*(S3), S366–S377. DOI: 10.1086/341859

Wright, C. D. (2012). Mechanistic explanation without the ontic conception. *European Journal for Philosophy of Science*, *2*(3), 375–394. DOI: 10.1007/s13194-012-0048-8

Ylikoski, P. (2012). Micro, macro, and mechanisms. In: Kincaid, H. (ed.) *The Oxford Handbook of Philosophy of the Social Sciences* (pp. 21–45) Oxford: Oxford University Press.

Ylikoski, P., & Aydinonat, N. E. (2014). Understanding with theoretical models. *Journal of Economic Methodology*, *21*(1), 19–36. DOI: 10.1080/1350178X.2014.886470

6 Interventions and manipulability

In economics and other policy-oriented disciplines, it is natural to identify causal claims with the notion of manipulability. However, the connection between causality and manipulability seems to also arise from everyday experience. As Daniel Hausman and James Woodward (2004, p. 856) admitted,

> [p]eople do not expect spontaneous correlations, and they do expect that there will be systematic relationships between (in)dependence relationships when intervention variables are off and when they are on, so that they can use each kind of information to learn about the other.

This explains why the manipulationist view on causality is popular among the philosophers of economics (Henschen 2018; Hoover 2001) and other practice-oriented sciences. This concept, roughly speaking, states that *X* causes *Y* if and only if a change in (intervention on) *X* modifies *Y*. Also, the commonsense view on causality identifies such relations with the possibility of influencing an effect by influencing its cause what instantiates Cook and Campbell's (1979) claim that "[t]he paradigmatic assertion in causal relationships is that manipulation of a cause will result in the manipulation of the effect" (p. 36). As James Woodward (2012, p. 234), a contemporary advocate of this stance, put it, the agency and interventionist theories have in common accepting the following definition of causality: "if *C* causes *E*, then if *C* were to be manipulated in the right way, there would be an associated change in *E*." This approach dates back to the early twentieth century, when Robin Collingwood (2001 [1940]) opposed the reductionist philosophy of logical positivism that has identified causal relations with empirical regularities (cf. Chapter 2) and grounded his metaphysical concept of causation in manipulability. The most notable supporters of the manipulationist view on causality include Menzies and Price's (1993) agency theory and Woodward's (2005 [2003]) interventionist stance.

This chapter is structured as follows. In Section 6.1, I discuss the philosophical theories of causality that belong to the manipulationist approach, consider their limitations, and review the philosophy of economics debates on the use of experimental and quasi-experimental research designs and the meaning of causality. In Section 6.2, I analyze cases of contemporary economic research that arguably presuppose different versions of the manipulationist notion. Doyle's (2007) instrumental variable (IV) estimation of the effects of foster care on children's economic outcomes and Pop-Eleches' (2006) natural experiment of the effect of abortion ban on socioeconomic success instantiate quasi-experimental research designs. The examples experimental studies are Dupas and Robinson's (2013) randomized field experiment of the determinants of a low rate of savings among the poor and Hussam et al.'s (2008) laboratory experiment studying bubbles on financial markets. Doyle's (2007) instrumental variable (IV) estimation of the effects of foster care on children's economic outcomes and Pop-Eleches' (2006) natural experiment of the effect of abortion ban on socioeconomic success instantiate quasi-experimental research designs. In Section 6.3, I focus on the use of the manipulationist evidence for policymaking and offer a solution to the extrapolator's circle based on the distinction between extrapolating from a random sample of a population to that population (populational extrapolation) and from a sample of one population to another population (extra-populational extrapolation).

6.1 The manipulationist theories and philosophical problems of experimentation

This chapter focuses on reviewing the philosophical debates related to the manipulationist approach to causality. In Section 6.1.1, I differentiate between agency and interventionist theories, argue that Woodward's interventionist definition can only be interpreted as an ontological stance because the notion of intervention is defined too strict to allow for testing. In Section 6.1.2, I study the philosophy of economics literature discussing definitions of causality related to the manipulationist notion and methodological problems of experimental and quasi-experimental research designs.

6.1.1 Manipulationist theories and their pitfalls

The label 'manipulationist approach' on causality covers two distinct views. On the one hand, agency theories define causality in terms of free (human) action influencing an effect. On the other, the notion of manipulability is considered theoretically and includes all possible interventions. Roughly speaking, both groups of theories accept the identification of causality with manipulability: X causes Y if and only if a possible or actual change in X causes a change in Y. Even though the two philosophical approaches to disentangling the notions of causality and manipulability coexist in contemporary literature, the agency theories dominated the earlier discussions, while the interventionist stance is more popular nowadays. Let me review the two families of manipulationist theories in chronological order.

Agency theories

The chronologically first discussion of causality in terms of manipulation-invariant relations is Robin Collingwood's (2001[1940]) disagreement with the then-mainstream reductionist stance of logical positivism. Taking causation as a metaphysical concept – i.e., one impossible to test – logical positivists advised getting rid of it from the endeavor of science and replacing it with the notion of laws understood as functional dependencies (cf. Chapter 2). Collingwood differentiated among three distinct notions of causation: (1) producing an outcome by human action, (2) singular causation between events, and (3) type-level causation resembling the regularity view (p. 285). According to this primary notion,

> that which is caused is the free and deliberate act of a conscious and responsible agent, and 'causing' him to do it means affording him a motive for doing it. For 'causing' we may substitute 'making,' 'inducing,' 'persuading,' 'urging,' 'forcing,' 'compelling' . . . .
>
> (p. 290)

Collingwood (1940 [2001]) argued that the first sense of the word 'cause' (i.e., causality as making or influencing by free human agents) is before the other notions (dependencies among events and the scientific events). This is the case because, according to his view, addressing the question of "'What is the cause of an event *y*?'" (p. 296), aims at delivering knowledge of how humans can bring about or prevent *y* from happening. Collingwood's (1940 [2001]) definition of causality is anthropocentric because, according to his view, "that which causes is something under human control, and this control serves as means whereby human beings can control that which is caused" (p. 286).

The entanglement of causal relations with human actions aimed at controlling the world is also present in Georg von Wright's (1971) theory of causality. Similar to Collingwood, Wright defines causality as relations that can be used by humans so that intervening on causes influences their effects. According to his view, "if *p* is a (sufficient) cause of *q*, then if I could produce *p* I could bring about *q*" (Wright 1971, p. 74). Given this passage, Wright's human-agency theory is deterministic: influencing a cause always brings about its effect. Such a notion is problematic when one considers probabilistic relations. In Wright's theory, the agency (manipulability) is also a more primitive notion, and causality can be reduced to it. In other words, relations are causal because they allow for using them to intervene. A more recent account of Peter Menzies and Huw Price (1993) also aims at reducing causal relations to manipulability by agents: "the agency theory correctly portrays causation as something analogous to a secondary quality-*as* a secondary quality" (p. 189, emphasis in original). This reduction is also motivated by the belief that humans from their early life experience intervene and change states of the world, but they cannot observe what it means to 'cause' an effect. However, contrary to Wright's theory, their approach merges the agency theory with the probabilistic account and allows for including in the set of causal relations such dependencies that do are not deterministic; i.e., only allowing for changing the probability that an effect occurs by influencing its cause.

According to Menzies and Price's (1993, p. 187) definition: "an event A is a cause of a distinct event B just in case bringing about the occurrence of A is an effective mean by which a free agent could bring about the occurrence of B." The probabilistic version of this view on causality uses the notion of agent's probability $P_A(B)$ that denotes the probability of event B in the case if A is brought about (p. 190). Such a definition of causality is a token-level definition (i.e., a definition of a causal relation between events). However, Daniel Hausman (1998, p. 86) indicated that the agency theory of causality could also be formulated in a way that allows for discussion of type-level relations. As I argue ahead, such generalization of the probabilistic version of Menzies and Price's (1993) agency theory makes it a good candidate for the definition of causality presupposed by economic studies aimed at uncovering relations invariant under interventions. This is especially likely considering the motivation driving Menzies and Price (1993) to formulate their version of the agency theories; that is, a need to oppose reducing causal relations to empirical regularities. As they argue, "[e]mpiricists need to keep in mind that human subjects have access to the world in two ways, as observers, certainly, but also as *agents*, capable of intervening in the processes of the world at will" (Menzies and Price 1993, p. 191, emphasis in original).

The main accusation raised against the agency theories is that they identify causal relations with human actions. Given that causal relations are reduced to human agency, it follows that there are no causal relations without human action, which is a problematic stance (cf. Hausman 1998). According to one version of this counterargument, similar to the question of whether chairs are chairs even if no one perceives them as this type of furniture, one can ask whether causal relations would disappear if humans were not present or if we were not able to act in the world as we are able now. According to another formulation of this counterargument, the reduction of causal relations to manipulability by humans leaves many (purportedly) causal relations unnoticed. For example, accepting a simple version of the agency theory leads to the conclusion that earthquakes are not caused by movements of continental plates because the plates are not manipulable. To save their own account from such accusations, Menzies and Price (1993) reformulated the notion of agency (manipulation) so that it covers both actually conceivable and theoretically possible manipulations that could be conducted if agents had other capacities. However, the problem that ocean tides cannot be considered as effects of the moon's movements (and with other naturally occurring dependencies) if manipulations are identified with agents' capabilities or theoretically possible interventions is especially sound for the natural sciences. Social sciences and, particularly economics, seem to be concerned with such causal factors that are, at least in principle, manipulable by humans.

Interventionist theories

The philosophers unconvinced by this defense put forward interventionist theories that refrain from reducing manipulations to agency and human ability to act and cause things in the world. Instead, those philosophers define manipulations as isolated interventions that are possible at least theoretically (i.e., while they may not be 'doable' by human agents, interventions should not be in disagreement with how the world works). The second noticeable difference is that the interventionist theories of causality neither focus on reducing causal relations to human actions nor consider the relation of manipulability as being ontologically superior (primary) to causality. Their purpose is to analyze the entanglement of the notions of causality and manipulability. Daniel Hausman and James Woodward (1999, p. 533) characterize roughly the interventionist view on causality as follows:

> [i]f *X* causes *Y*, one can wiggle *Y* by wiggling *X*, while when one wiggles[1] *Y*, *X* remains unchanged. If *X* and *Y* are related only as effects of a common cause *C*, then neither changes when one intervenes and sets the value of the other but both can be changed by manipulating *C*.

According to Nancy Cartwright (2003, p. 215), "we are not, after all, interested in invariance itself but pursue it as a test for causality," and therefore studying the conditions under which relations are invariant under interventions allows for obtaining knowledge on causality. Ahead, I focus on analyzing the notions of intervention and invariance employed in discussions of the Bayesian nets[2] and James Woodward's (2005 [2003]) theory.

DAGs (or directed acyclic graphs), when used in the context of causal inference, refer to a set of algorithmic methods based on Bayes' theorem developed by Judea Pearl with his collaborators (Pearl 2000; Pearl and Verma 1995; Sprites et al. 1993). These algorithms infer causal structure by means of analyzing probability distributions of variables. The results are represented in the form of acyclic graphs. If one rejects the agency theories, one needs to define interventions in a way that avoids the difficulties raised by the opponents of defining causality in relation to human actions.

Judea Pearl (2009, p. 71) employed the notion of 'atomic' intervention. Such interventions are atomic because they only modify the target variable. Specifically, according to Pearl, interventions change the probability distribution of a targeted variable so that the variable that is the target of intervention is screened off from their parents (i.e., the variables located before it in a directed graph). In agreement with Pearl's views, Chris Meek and Clark Glymour (1994, p. 1010) defined interventions as a replacement of probability distribution (resulting from the determinants of a variable) of the variables directly influenced by the intervention. Such a replacement of probability distribution can be labeled 'arrow breaking (cf. Sprites et al. 1993) in the sense that the arrows in a directed acyclic graph that connect the targeted variable with its causes disappear by intervention. Interventions are required to be conducted in a way that ascertains that the targeted variable is not related (statistically) to the variables other than its effects. As Nancy Cartwright (2002, p. 438) put it, "*X* should be produced by a method for which the resulting *X* values are probabilistically independent of any other quantities that are not effects of *X*" (i.e., independent of its causes and all other variables in a graph with the exception of its effects). The definition of atomic intervention, despite being possibly an inadequate description of the actual interventions in economic policymaking, overlaps with the notion of causal Markov condition that is crucial for causal inference with DAGs (Hausman and Woodward 1999). While such requirements on interventions ascertain mathematical tractability and justify causal conclusions obtained by means of Bayesian nets, they may be rarely met in economics. For instance, fiscal stimulus interventions are likely to be negatively correlated to the pace of economic growth. This may explain why economists are quite skeptical of Bayesian nets that are rarely used in this discipline.

Furthermore, only some interventions fulfill the definition of the 'setting' intervention (set $X = x$); i.e., determine the value of a targeted variable or its probability distribution. Other interventions only influence the value of a targeted variable, but are unable to screen it off from their antecedents in a graph. For instance, a policymaker can raise safety requirements for mortgages but is unable to screen the value of housing credit from other determinants (such as the demand for houses).

Moreover, Bayesian nets seem to be susceptible to the same criticism as econometric atheoretical modeling (cf. Chapter 3): excluding the possibility that results are driven by a common cause (which is not included in a dataset under study) is in principle impossible if one uses observational data. As Hausman and Woodward (1999, p. 560, emphasis in original) put it, "if the joint probability distribution over the variables in the graph was generated by deterministic processes that include such 'wiggling,' then **CM** [causal Markov condition] should hold." However, if one uses purely observational data, then the causal Markov condition (and the DAGs) can uncover spurious causal relations (cf. Cartwright 2001; Dawid 2010). The possibility of common-cause fallacy, which is particularly striking in the realm of economics due to the presence of the phenomenon of multicollinearity, makes randomized studies aiming at equal distribution of all confounding factors a better tool for causal inference in comparison to econometric modeling and Bayesian nets. As Cartwright (2003, pp. 214–215) put it, "[r]andomized treatments/control experiments are the gold standard for establishing causal laws in areas where we do not have sufficient knowledge to control confounding factors directly."

Despite these objections, the Bayesian nets literature obtains interesting results for the relation between the accuracy of causal (statistical) models and warranting invariance under intervention. For instance, Cartwright (2003, p. 208) proved that the correctness of a causal linear structure implies being invariant under intervention (in the sense that interventions do not influence the structure understood as causal laws). While this conclusion holds only under relatively strong assumptions (linearity and determinism of a system) that are rarely meet in research practice,[3] it can be intuitively employed to argue that testing for invariance under intervention can deliver strong evidence for the causal nature of relation under consideration. Similarly, Hausman and Woodward (1999) argued that the equations that represent causal relations are true if they are invariant under interventions modifying the values of independent variables (level invariance).

WOODWARD'S THEORY

While the philosophical and statistical literature on Bayesian nets has never entered mainstream economics and philosophy of economics directly, it has inspired James Woodward's interventionist theory that (with slight modifications) has been considered as a definition of causality adequate to macroeconomics (Henschen 2018). James Woodward's (2005[2003]) theory emerges from the observation that different philosophical perspectives on causality (such as previous manipulationist accounts, counterfactual theories, and the discussions of invariance within the DAG frameworks and among econometricians) have been developed in separation. Woodward's theory of causation can be considered as an attempt at a reconciliation of these approaches. Woodward modified agency theories by redefining the notion of intervention (imported from the Bayesian nets) so that it no longer depends on human agency and uses counterfactuals to define causal relations. Also, the view that changes in the values of variables are the relata of causal claims (Woodward 2005 [2003], p. 112) indicates the inspirations from these methods of statistical inference. However, his account also differs in relevant aspects from these approaches.

For example, despite employing counterfactual claims, Woodward (2005 [2003], p. 15 et seq.) formulated his interventionist theory using manipulationist counterfactuals, contrary to Lewis' (1973) earlier views (cf. Chapter 4). The difference can be exemplified with the relation between barometer readings and weather, which is often considered by philosophers. Within Lewis' framework, the counterfactual describing the dependency (e.g., 'if the barometer's indicator showed lower pressure, there would be a storm') is true. On the contrary, the use of manipulationist counterfactuals by Woodward (e.g., 'if the barometer's indicator were *lowered*, there would be a storm') makes the relation spurious. Woodward (2016) argued that the usual formulations of the manipulationist theory of causality are problematic because they imply that relata of causal relations are causes and effects only under interventions. In other words, there are no causal relations without manipulations. This view seems to be a counterintuitive implication of viewing causality as relations invariant under interventions. To solve this problem, Woodward put forth the counterfactual formulation of the manipulationist account according to which "*A* causes *B* if and only if *B* would change if an appropriate manipulation on *A* were to be carried out" (p. 21).

The motivation for putting forward a manipulationist theory is Woodward's (2012) belief that it delivers a natural distinction between causal and "purely correlational" (random) associations. If an inferred causal relation is spurious, an attempt at influencing an effect by manipulating its cause is doomed to failure. Daniel Hausman and James Woodward (2004, p. 847) described this view as follows:

> [w]hen *X* and *Y* are correlated and *X* does not cause *Y*, one expects that when one manipulates *X*, the correlation will break down. By contrast, if *X* causes *Y*, one expects that for some range of values of *X*, if one is able to manipulate those values, one can thereby control the value of *Y*.

At first glance, the manipulationist view on causality describes the types of relations that can be uncovered by experimental research. However, Woodward (2005 [2003]) himself admitted that interventions are "idealized experimental manipulation[s] carried out on some variable *X* for the purpose of ascertaining whether changes in *X* are causally related to changes in some other variable *Y*" (p. 94). The manipulations do not need to be doable by humans.

Furthermore, the strict conditions for interventions specified by Woodward (2005 [2003], p. 98) make implementing such interventions difficult or even impossible. Therefore, Woodward's theory is of ontological nature: it studies the connection between manipulability and causality, but does not limit causal relations only to those having actually manipulable causes. Comparing the notion of intervention (considered ahead) to idealized experiments does not exclude other types of evidence. Woodward (2005 [2003]) accepts even observational evidence (given that the influence of confounders is controlled for).

Woodward's theory belongs to the tradition that rejects the attempts at reducing causality to the agency on the grounds that they raise the problems of anthropocentricity and circularity. While solving some problems, this move raises the problem of delivering a detailed definition of intervention. Woodward (2005 [2003]) and Daniel Hausman (1998) agree that such a definition should require interventions to be 'surgical' so that it influences only the target of manipulation and no other aspect of reality. As Woodward (2005 [2003], p. 130) put it, intervention should be "sufficiently fine-grained and surgical that it does not have any other effects on the tides besides those" planned. The surgical nature of interventions means that they influence only the target of intervention and no other variable. Furthermore, interventions should be causally independent of their target system. They, according to Woodward, should occur spontaneously; i.e., the variable denoting intervention (*I*) should be statistically independent of all variables in the set *V*, with the exception of its effects.

Woodward (2005 [2003]) employed the view on interventions from the Bayesian nets literature that uses the notion of 'setting' interventions. In the intuitive reading,

> the introduction of the intervention . . . must 'break' this existing causal connections between *V* [the set of all variables] and *T* [treatment variable]

> so that the value of T is now set entirely exogenously by I and is no longer influenced by V.
>
> (p. 96)

Formally, Woodward (2005 [2003], p. 98) puts forward the following four requirements. First, I is an intervention that modifies X if it directly influences X. Second, causal interventions should prevent other causal mechanisms from operating (i.e., I should act as a switch). Third, interventions should not have a direct influence on the targeted variable Y. Fourth, interventions have to be statistically independent of any determinants of the target of intervention and its effects.

The second requirement, that interventions should act as a switch – i.e., under intervention, "X ceases to depend on the values of other variables that cause X and instead depends only on the value taken by I" (Woodward 2005 [2003], p. 98) – may be problematic if one wants to interpret Woodward's definition of causality as an epistemic stance (i.e., a description of a test for causality). This condition requires that interventions should break other causal dependencies so that only the variable on which the intervention was conducted is a cause of the target of intervention (Woodward 2005 [2003], p. 97). However, in the realm of social sciences, setting interventions are rarely accessible to scientists or policymakers. Furthermore, interventions are likely to target more than one mechanism. This contradicts Hausman and Woodward's (1999, p. 542) view that "[i]nterventions that set the value of a single variable thus disrupt only the single mechanism that previously determined the value of that variable." My doubt will be more visible after I discuss case studies in the following Section 6.2.

After defining the notion of intervention within Woodward's framework, I can now move to introducing his definition of direct causality. According to Woodward (2005 [2003]),

> [a] necessary and sufficient condition for X to be a direct cause of Y with respect to some variable set V is that there be a possible intervention I on X that will change Y (or the probability distribution of Y when all other variables in V besides X and Y are held fixed at some value by additional interventions that are independent of I.
>
> (p. 55)

In his definition of direct causality, Woodward (2005 [2003]) uses the notion of *possible* intervention. In the case of agency theories, this notion is related to human capabilities. However, given that Woodward refrained from defining interventions in relation to human action, what it means that an intervention is possible needs to be reconsidered. Woodward (2005 [2003]) argued for a broad view as follows:

> if 'possible' is taken to mean something like 'within the present technological powers of human beings,' then NC [the necessary condition formulation

> of the definition of causality] has the obviously unacceptable consequence that *X* cannot cause *Y* when human beings lack the power to manipulate *X*. (p. 46)

On this basis, possible interventions are such that they agree with laws of nature but do not have to be doable by humans.

One of the primary purposes of Woodward's account is to offer a way of discriminating between accidental and law-like regularities (2005 [2003], p. 240). According to his stance, only those regularities are stable (are not accidental) that are invariant (to a certain degree) under interventions. Woodward requires only a moderate version of invariance under interventions. As he put it, "[i]nvariance, as I understand it, does not require exact or literal truth; I count a generalization as invariant or stable across certain changes if it holds up to some appropriate level of approximation across those changes" (p. 236). According to Woodward (2005 [2003], pp. 315–316), a linear regression from a structural-equation model (e.g., $Y = \hat{a}X + \varepsilon$) can[4] be interpreted causally in agreement with his interventionist theory under the *ceteris paribus* clause. In such case: if the intervention $X = x$ were conducted, then *ceteris paribus* $Y = \hat{a}x$. While I have argued that structural equation modeling, as practiced in econometrics, presupposes a version of the regularity view on causality, it is true that if (1) a model represents the actual causal structure, and (2) the values of other variables stay unchanged during intervention, then the interventionist reading is justified. However, these two conditions are rarely met in practice, and even if they were, policymakers would be unable to decide whether it is the case.

It is fruitful for further discussions to mention Nancy Cartwright's (2007b, p. 136) differentiation between testability by experiments and manipulability theories. Cartwright (2006, 2007a) distinguished between manipulability and testability in the following way. The requirement of manipulability states that interventions influence causes, but the influence can be unobservable due to the noise created by other causes of the phenomena under consideration. On the contrary, the requirement of testability "requires that there exists an 'intervener'/'manipulation' for every factor, not just for causes" and "manipulating a cause changes its effects; but also manipulating non-causes of a factor does not change it" (Cartwright 2006, p. 202). Even though Cartwright includes the views of Hausman and Woodward (2004) in the testability camp on the basis that interventions need to produce *ceteris paribus* observable changes, I think that Woodward's (2005 [2003]) theory of causality does not allow for a direct test of his definition. This is the case because Woodward's definition is too strict, and some conditions cannot be met in practice: it describes an ideal rather than an actual intervention (cf. Reiss 2019, Section 6.2).

Apart from Woodward's interventionist theory, more practical views are present in the literature. For example, Cook et al. (2002) argued in their notable book that the interventionist approach to causality is presupposed by the scientists who ground their causal inferences in experimental and (in case of the lack of possibility to conduct them) quasi-experimental research methods. In such cases, causes are required to change, on average, their effects.

Reutlinger (2012) opposed Woodward (2005 [2003]) on the grounds that interventions should be possible to conduct to allow for testing. Furthermore, Woodward's interventionist theory has received criticism similar to the voices opposing agency theories. For example, Michael Strevens (2007, 2008) accused his interventionist account of being circular. Woodward's response, which may be unconvincing for those philosophers who believe that any meaningful account of causality needs to reduce such relation to some other notion, is that the use of invariance under intervention, which is a feature of causal relations, in explanandum does not necessarily make definition useless. In other words, Woodward admits that there is some ground for accusing his account of circularity, but defends it for being informative in disentangling the notion of causality and manipulability. This defense seems to be accepted among the philosophers of economics, since Woodward's account has been considered as an account adequate to macroeconomic causality (Henschen 2018) and resembles Kevin Hoover's views (see Hoover 2001).

6.1.2 The manipulability account in economic methodology

In this subsection, I focus on reviewing the philosophy of economics literature related to the manipulationist approach to causality. First, I discuss the voices supporting a version of the manipulationist approach to causality. These are Tobias Henschen's recent argument supporting his version of the manipulationist notion, Kevin Hoover's approach to inferring the direction of causal relations, and David Hendry's discussion of superexogeneity. Second, I study methodological discussions of experimental and quasi-experimental research designs. In the following Section 6.2, I exemplify the research methods presupposing a version of the manipulationist view on causality with case studies. In Section 6.3, I discuss the use of manipulationist evidence for policymaking.

Manipulationist causality in philosophy of economics

Recently, Tobias Henschen (2018) has supported a version of the manipulationist definition of causality as a concept adequate to macroeconomics. The vagueness of the notion of 'adequacy' in his paper has resulted in an exchange of arguments on whether a chosen manipulationist definition is sufficiently broad to describe causal relations represented by all macroeconomic causal models. I and Robert Mróz (Maziarz and Mróz 2019, 2020) argued that while it is adequate to some macroeconomic models, other macroeconomic models (VAR models standing for probabilistic dependencies and DSGE models representing mechanisms) can be considered to be causal only if one accepts other definitions of causality. Given that economists refer to such models as causal models and their direct application to macroeconomic policymaking, a version of moderate causal pluralism seems to be an adequate view on causal relations represented by macroeconomic causal models. Henschen (2020) responded that his manipulationist definition is normative in nature, and therefore, it should serve the purpose of differentiating between causal and noncausal models. While I agree in this book that having evidence that a relation under study is invariant under intervention is superior to other types of evidence, I do not think that other notions of causality (or types of evidence presupposing these different notions) should be excluded from economics (cf. Maziarz and Mróz 2020).

Although the discussion is far from being settled, Henschen's definition accurately describes the relations represented by at least a group of macroeconomic models. Let me first introduce Henschen's definition and later review other voices supporting manipulationist definitions of causality present in the methodological literature. The definition states that "*X* directly type-level causes *Y* if and only if there is a possible intervention on *X* that changes *Y* (or its probability distribution) while all causal parents of *Y* except *X* remain fixed by intervention"[5] (Henschen 2018, p. 16). For Henschen (2018), interventions set the value of targeted variables, influence *X* (targeted variables) and *Y* only through paths that go through *X*, and are statistically independent of any other variable in the set. Finally, interventions *I* can be either modification of parameters or variables. This definition results from Henschen's study of Woodward's theory of causality, Kevin Hoover's (2001) book and Joshua Angrist and Guido Kuersteiner's (2011) potential outcome approach.

Henschen (2018) has motivated his study by indicating that macroeconomists rarely explicitly define causal relations and points out the exceptions of Kevin Hoover and David Hendry. Other econometricians concerned with methodology of causal inference also support a version of the manipulationist definition of causality. For example, James Heckman (2008) admitted that "[e]conomists focus on causality from the perspective of policy evaluation" (p. 5). Interestingly, Henschen (2018) refrained from discussing the views (see Hendry 2000), who identified causality with superexogeneity, given that this conception clearly belongs to the manipulationist approach (cf. Hoover 2008).

Robert Engle et al. (1983) differentiated between and formalized the concepts of weak exogeneity, exogeneity, and superexogeneity put forward by Tjalling Koopmans (1950). According to their proposition, a variable z is weakly exogenous if the inference of a set of parameters λ in a model conditional on z does not involve a loss of information (p. 278). Strongly exogenous are such variables that are weakly exogenous and are not Granger-caused by any other variable from a model. Finally, superexogeneity denotes the situation when parameters λ in a model conditional on z are stable when "mechanism generating z changes. Such changes could come about for a variety of reasons; one of the most interesting is the attempt by one agent to control the behavior of another [. . . (notation slightly modified)]" (Engle et al. 1983, p. 278). In other words, superexogeneous variables can be intervened on without influencing causal structure (i.e., model parameters). While this idea resembles Woodward's theory, the superexogeneity definition is different in regard to the view on what are interventions. For Hendry and other econometricians, interventions alter the values of targeted variables rather than setting them.

Another issue considered in the literature is to indicate what an intervention influences in a theoretical framework. According to Pearl (2009, 2000), and Pearl et al. (2016), who work in the graph-theoretic paradigm, each equation (within a structural equation model) represents a distinct causal mechanism[6] and interventions break the operation of such a mechanism and sets the value of a target variable. In other words, the value of variable *X* is entirely determined by intervention. Kevin Hoover (2001) concluded his study of research on the relation between money supply and price level, indicating that it is more fruitful to state that interventions change the values of parameters instead of the values of variables (cf. Hoover 2012, p. 103; Meyer 1995). Hendry and other econometricians working on superexogeneity tests define interventions as changes in the values of variables so that *I* is added to the value of *X* at time *t* of intervention instead of screening of *X* from its determinants and setting its value independently. This opposes Woodward's (2005 [2003]) theory employing the concept of 'setting' interventions from the DAG literature. Also for Stephen LeRoy (2016, 2018), who has studied structural equation modeling, interventions are modeled as changes in the values of variables instead of changes of parameters. Stephen LeRoy (2016, 2018) studied intervention-invariance from a different angle and addressed the question of the features that allow for using a structural equation model for intervening: a cause cannot influence an effect both directly and indirectly because, in such a case, the strength of causal relation depends on the size of intervention and hence the relation under consideration lacks intervention neutrality.

Kevin Hoover is another notable econometrician and philosopher who has worked extensively on causal inference. Hoover's (2012, p. 91) overview of the debates on causal inference in philosophy of economics distinguishes between a reductive approach to considerations focused on the ontology of causality and analyzing the methods of causal inference that are based on the presupposition that "[c]ausality is related to invariance" (Hoover 1990, p. 224). Hoover (2011, p. 342) seems to accept a modified version of Woodward's manipulability account. According to Hoover, one of the features of causality (possibly, it is the crucial feature) is the possibility of influencing an effect by manipulating its cause (Hoover 2001). However, Hoover's work is concerned with delivering epistemic guidance instead of an ontological theory reducing causality to intervention-neutrality. The resemblance of Hoover's and Woodward's views is visible in that both accept the use of counterfactuals for analyzing causality. As Hoover (2011) put it,

> counterfactuals play an essential role . . . in the notion that the causal relationship is to be evaluated in isolation by holding other variables fixed. There may be no way actually to achieve such holding fixed and, like Lewis, Woodward is willing to countenance the semantic device of 'small miracles' to achieve the necessary isolation.
>
> (p. 341)

Henschen (2018, p. 10) extracted from Hoover's work the following two definitions and indicated that the latter should be preferred. First, "X directly type-level causes Y if and only if Simon's hierarchy condition and Simon's condition of privileged parameterization hold." Second, "X directly type-level causes Y if and only if Simon's hierarchy condition and Hoover's parameter-nesting condition hold." According to Hoover (1990, p. 215), the manipulationist view on causality in economics dates back to Herbert Simon's (1957 [1977]) identification of causal relations with controllability: "A causes B if control of A yields control over B" (Hoover 1990, p. 215). Both Hoover and Simon take model parameters as the relata of causal claims. Contrary to Simon and Hoover, Cooley et al. (1984) argued that all parameters should be constant in every model, and interventions should be regarded as a change in the value of a variable. The latter view on the relata of causal claims seems to be more popular among econometricians.

Econometric inference of intervention neutrality

However, apart from his conceptual work, Hoover also developed a method of causal inference known as Hoover's test. In *Causality in Macroeconomics*, Hoover (2001) strived for putting forth methods of causal inference grounded in the interventionist approach as a third way distinguished from the Cowles Commission tradition of estimating econometric models heavily dependent on theory (cf. Chapter 2) and the atheoretical econometrics (cf. Chapter 3) grounded in the study of time precedence. The problem with econometric models is that correlation is a symmetrical relation. Therefore, data alone are grounds for a set of inconsistent models. The choice among models postulating inconsistent causal structures can be based on aprioristic theory, knowledge about time precedence, and, according to Hoover (2001) and Herbert Simon (1951), extra-theoretical knowledge about time of interventions. The use of Hoover's test can be explained as follows. For simplicity, assume that we have a two-dimensional macroeconometric model. In such a situation, there are two possible model specifications that support two contradictory causal hypotheses. Accepting the convention that causes are located on the right-hand side of the equation, Model 1 suggests that $B \rightarrow A$. Model 2 is, in contrary, evidence that $A \rightarrow B$.

1 $A = C_1 + \alpha_1 \star B + \varepsilon_1$

2 $B = C_2 + \alpha_2 \star A + \varepsilon_2$

Hoover (2001, ch. 8–10, 2012) argued that the use of knowledge about interventions could be employed to solve the problem of underdetermination. In the case of the two regressions that are observationally equivalent, one can estimate a regime-switching model that would entail the (non-data-driven) moment of a change in economic policymaking and, based on this, appraise the true causal structure (cf. Henschen 2018, pp. 13–15 for a detailed discussion). A similar solution has been offered by Simon (1957, 1977, pp. 63–65) in response to the high reliance of the Cowles Commission methodology on economic theorizing (cf. Chapter 2). In detail, Simon advised considering causal orderings of each plausible set of structural equations in order to choose the model that does not imply theoretically impossible mechanisms of interventions:

> in order to discuss causality, a second language (a 'metalanguage') describing the relationship between the 'experimenters' and the model [is needed]. The terms 'direct control' and 'indirect control' are in this metalanguage. Thus, in our metalanguage we have an asymmetrical relationships (>) – behavior of experimenters > equation coefficients > values of variables – that must be introduced in order to establish the asymmetrical causal relationship (→).
>
> (Simon 1977, p. 65)

Simon's (1951) discussion of whether being under the direct control and the production of wheat under indirect control indicates that his notion of controllability resembles the views on interventions present in interventionist theories in comparison to the agency approach. Such interventions should be theoretically possible, considering theoretical knowledge about mechanism connecting variables, but they do not need to be within human capacities.

A method similar to Simon's and Hoover's approach in using the knowledge on the timing of interventions but based on the concept of superexogeneity was developed by Engle and Hendry (1993). The test of superexogeneity studies whether model parameters and variance of error terms are stable in time. Such a test presupposes that econometric models are estimated on datasets that cover both periods when no intervention occurred and periods of policy changes (cf. Figure 6.1). If an econometric model depicted a spurious correlation, then, according to Engle and Hendry (1993), interventions, by changing causal structure, would influence model parameters and/or variance of error terms during active policymaking. Therefore, studying the stability of parameters by estimating model on subsamples allow for concluding whether relations are invariant under interventions, and, at least in principle, differentiate between econometric models representing actual causal structures from those depicting spurious correlations.

In practice, while testing for superexogeneity keeps down the risk of common-cause fallacy, it does not exclude it since such a test for a model of two variables having a common cause but themselves causing different variables would lead to a false-positive result. For instance, if the relation between X and Y (tested for superexogeneity (Figure 6.2, dotted line) has a common cause A (or common causes A_1; A_2; A_n), but they affect the values of variables that are

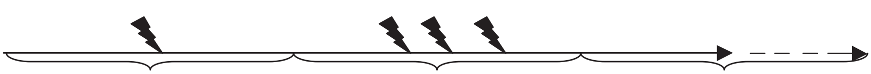

Figure 6.1 Different numbers of interventions (symbolized by lightning) in subsamples allow for estimating whether model parameters are stable under interventions

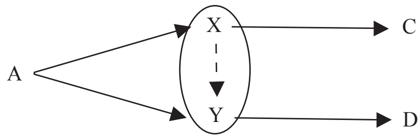

Figure 6.2 A (simplified) causal structure leading to false results of superexogeneity tests. The variables included in the model under test are in the circle

not included in a model (*C*; *D*), then external shocks (interventions) on *X* are unlikely to result in observable changes in parameter values. In such a situation, an intervention on *X* changes *C* but has no influence on *Y*. If the causal structure does not change during the period under study (i.e., *X* and *Y* have a common cause), then superexogeneity tests show that parameter values stay constant and deliver false evidence for the hypothesis that the relation between *X* and *Y* is invariant under intervention.

Natural experiments and quasi-experimental research designs

Both Simon-Hoover tests and methods based on the notion of superexogeneity have been developed as an aid for time-series econometrics and, despite being related to the concept of intervention, belong to the repertoire of observational studies. However, econometricians and statisticians working in other disciplines, having in mind the gold standard[7] of a randomized experiment,[8] have developed research designs that mirror this superior approach to causal inference. Thad Dunning (2012, p. 1) admitted that natural experiments entered scientific practice in the 1980s, but their number has skyrocketed in the first decade of the twenty-first century. Because of the novelty of this approach to causal inference in economics, the labels present in the literature seem to be ambiguous to some degree. For example, Julian Reiss (2016, p. 138) defined natural experiments as a research design including two populations (receiving treatment and control group). According to his view, the use of instrumental variables (IV) estimation that artificially divides the sample into treatment and control group (and which I consider ahead) to infer causal claims can also be considered as a natural experiment given the quasi-random assignment. Similarly, James Heckman (2008, p. 51) characterized this method of causal inference as searching "for credible sources of identifying information for causal parameters, using ideal random experiments as a benchmark."

Others take 'natural experiment' to be a label for the situation in which an intervention is not controlled by an experimenter, but, to the contrary, happens independently (is created by nature). As Breed Meyer (1995, p. 151) highlighted, "[g]ood natural experiments are studies in which there is a transparent exogenous source of variation in the explanatory variables that determine the treatment assignment." Thud Dunning (2012, p. 12) characterized the crucial feature of natural experiments indicating that "in a valid natural experiment, we should find that potential confounders are balanced across the treatment and control group, just as they would be in expectation in a true experiment." On the contrary, nonrandom assignments are the nature of quasi-experimental research designs (Achen 1986, p. 4). According to this distinction, which I employ, natural experiments make use of randomization (or quasi-randomization) by factors located beyond the control of the researcher, while in quasi-experimental research designs, it is the role of the researcher to construct treatment and control groups. However, still other distinctions and definitions are present in the literature. For instance, Jaasjet Sekhon (2009) differentiated between natural experiments and regression discontinuity design. Using either Sekhon's or Achen's and Dunning's distinctions leads to the conclusion that regression discontinuity design, which is often used in macroeconomic 'natural experiments' (e.g., Lalive 2007; Caughey and Sekhon 2011) is actually a quasi-experimental approach. This is the case since the use of time-series data and constructing treatment and control groups from data from the periods of just before and after interventions lead to obtaining groups that are confounded by time.

The flourishing of design-based studies that we can observe today seems to have been started by Edward Leamer's (1983, p. 31) criticism of atheoretical econometrics. Despite later debunking this view, he admitted that "[r]andomization seems to be the answer" to the methodological difficulties connected to observational studies. The rationale for using natural experiments or quasi-experimental research designs is the epistemic, financial, etc., impossibility (or impracticality) of conducting experimental studies. As Joshua Angrist and Jurn-Steffen Pischke (2010) put it,

> experiments are time consuming, expensive, and may not always be practical. It's difficult to imagine a randomized trial to evaluate the effect of immigrants on the economy of the host country. However, human intuitions or the forces of nature can step into the breach with informative natural or quasi-experiments.
>
> (p. 4)

Heckman (2000, p. 51) indicated that the popularity of the design-based approach is growing because "[i]t rejects the use of structural econometric models because, according to its adherents, such models do not produce credible estimates and impose arbitrary structure onto the data."

The quality of evidence from natural experiments and quasi-experimental research designs depends on the validity of studies. If the natural randomization or artificial quasi-random assignment does not warrant approximately equal assignments between the treatment and control groups, then the usual criticism of observational studies may apply. On the contrary, if these research designs are successful in mimicking randomized controlled trials, then the methodological appraisal of experimental studies applies. The main methodological criticism of the design-based approach to causal inference focuses on the fact that the research questions tackled in this way are determined by nature in the sense that natural experiments can be found only in some areas of economics. Furthermore, the questions addressed by these studies are narrow. According to some economists, the questions are too narrow and too localized to inform general policy questions or theoretical development. For example, Angus Deaton (2010) argued that too much emphasis on natural experiments might lead to establishing very narrow research programs that will be unable to enlighten policy practice. Another strain of literature highlights that natural or quasi-experimental research designs are, in fact, observational studies and not experiments. Therefore, marketing the design-based studies as (natural or quasi-) experiments may be an attempt at hiding the shortages of econometric techniques employed in processing data (Sims 2010, p. 59).

Dunning (2012, p. 18) distinguished between the following three different types of natural experiments: 'standard' natural experiments, regression-discontinuity designs, and instrumental-variables designs. The first type mentioned by Dunning delivers results most similar to actual experiments (i.e., RCTs). Standard natural experiments are such studies that analyze the effects of some intervention on the basis of analyzing the difference between treatment and control groups. The only difference from RCT is that group assignment and intervening lies beyond the control of the experimenter but – what is crucial for the validity of results – is random. The typical examples of such studies are designs based on lottery assignments to treatment and control groups. For example, Angrist and Krueger (1992) used the Vietnam War draft lottery (the American army randomly chose participants) to estimate the effects of education on income. Regression-discontinuity designs, which are a popular research design in macroeconomics, are based on analyzing time-series data and comparing observations from periods before and after interventions. However, regression-discontinuity designs can also be employed to cross-sectional data. In such cases, the analysis focuses on estimating the effect of intervention by comparing observations above (treatment group) and below (control group) a certain threshold. For instance, James Banks and Fabrizio Mazzonna (2012) evaluated the effects of education on old age cognitive abilities by comparing the difference between those slightly above the recruitment limit and those not accepted to college. In this case, the allocation to control and treatment groups are not random, and therefore the confounding needs to be corrected statistically. David Lee and Thomas Lemieux (2009) related the regression-discontinuity designs to the potential outcome approach. Ahead, I exemplify this research design with a study of the effects of an abortion ban in Romania on children's wellbeing (Pop-Eleches 2006).

A popular statistical technique used in quasi-experimental research designs is instrumental-variable (IV) estimation. It has been developed initially as an analytic method aimed at dealing with the problem of omitted variables. In such a case, standard econometric techniques bias the estimates (Hogan and Lancaster 2004). The IV method relies on the use of another, 'instrumental' variable that influences the 'intervention' variable, but not the target variable by way other than by the 'intervention' variable. In other words, using the IV approach, one should include into a regression an instrument (Z) that is correlated with X but is not correlated with Y (cf. Angrist and Krueger 2001, p. 73). If one employs the causal interpretation, then the above condition can be reformulated as follows: Z does not influence Y by way other than the intermediary of X: $Z \rightarrow X \rightarrow Y$. The instrumental variable (Z) is chosen on the grounds of aprioristic theoretical considerations, and not on the grounds of data itself.

The IV method can be used to analyze data generated by a natural experiment when a policy intervention or natural forces produced a situation similar to random experimentation. Angrist and Krueger (2001, p. 73) highlighted that "[r]ecent years have seen a resurgence in the use of instrumental variables in this way . . . to exploit situations where the forces of nature or government policy have conspired to produce an environment somewhat akin to a randomized experiment." However, I need to underline that, historically, the IV technique has been developed primarily as a method of dealing with unfulfilled assumptions of OLS estimation; in particular, to solve the problem of explanatory variables correlated to explained variable. Therefore, it is its use rather than the technique itself that allows for making conclusions about the effects of an intervention under study. The validity of manipulationist conclusions depends primarily on the choice of instrumental variable. The perfect instrument needs to be "uncorrelated with the omitted variables and the regression error" (Angrist and Krueger 2001, p. 73), but it also needs to be causally and not only statistically related. As Pearl (2000, pp. 136–137) highlighted, the causal definition of the results obtained by IV is justified if and only if the modeled equation is of 'structural' nature instead of being purely correlational. In a similar vein, Julian Reiss (2005) claimed that the IV method relies on reasonably strong assumptions that are usually not tested by econometricians. If studies using IV estimation are valid, then they allow for discovering relations invariant under interventions (cf. Freedman 2007). In the following Section 6.2.4, I analyze Doyle's (2007) study of the influence of foster care on economic outcomes in adulthood as an example of this type of quasi-experimental research design.

Experimental studies

Random-assignment experiments (randomized control trials, RCT) can arguably (cf. Section 6.2) be believed to instantiate a direct application of the manipulationist account. The rationale for this view is that if the treatment and control groups are randomized (each participant of an experiment has an equal chance of being chosen for each group), the confounders will be equally distributed among the two groups, at least if the number of participants were infinite. In the limit, the influence of all other factors should average out. Hence, the difference between the treatment group and control group shows the average effect of the intervention on the sample (average treatment effect) (cf. Cartwright 2007b).

Contrary to Milton Friedman's opinion voiced in his famous essay (1994 [1953]) that experimental testing is rarely possible in economics, it becomes more and more popular nowadays. For example, Christopher Starmer (1999) observed the rapid growth of experimental research in economics two decades ago. In a similar vein, Alvin Roth (1988), the author of one of the first handbooks on the method of laboratory experimentation in economics, has indicated that experiments are now a standard way of researching certain economic phenomena (p. 974).

The views on the role of experimental evidence within economics have changed. While earlier experimental studies were perceived as tests for economic theory, experimental economics is now taken as a stand-alone project capable of addressing policy questions (Wilde 1981; Starmer 1999). A representative example of the former stance is Vernon Smith's (1982) opinion that experiments "are sufficient to provide rigorous controlled tests of our ability as economists to model elementary behavior" (p. 935). On the contrary, Francesco Guala and Mittone (2005) argued that experimental results should not be judged from the perspective of mainstream theoretical economics, but they can be used as evidence addressing policy questions. As they put it,

> [t]he point is that their success should not be made a *precondition* for doing 'proper' economic science. . . . Experimental economists *are* doing proper economics right now, and can keep doing it without incorporating the conditions that specify a theory's application within the theory itself.
>
> (p. 189, emphasis in original)

Distinguishing between the following two types of experiments is useful for the discussion. First, randomized laboratory experiments aim at constructing an artificial setting isolated from other influences. To conduct such studies, economists usually use computer programs that simulate market interactions or economic games, but sometimes limit technological aids to simple questionnaires. Participants virtually always get monetary rewards for participation, and the experiment's value depends on their decisions. Two notable subgroups of laboratory experiments are virtual market experiments and game-theoretic experiments. Second, randomized field experiments are studies that aim at measuring the average influence of an intervention in the 'field'; i.e., in the economy. John List (2007, p. 9) distinguished among the following three types of field experiments. First, artifactual field experiments differ from laboratory experiments in refraining from using students as participants, but are also conducted with questionnaires or create artificial settings. Second, framed field experiments also use field context as a part of the experimental setup. Third, the most methodologically advanced, according to List's classification, are natural field experiments. These are conducted on participants located and acting in their natural environment. The popularity of field studies is rapidly growing. Even though both laboratory and field experiments proceed "by comparing the behaviour of two randomly selected groups under different experimental conditions" (Starmer 1999, p. 10), the two experimental designs may have different policy implications.

Generally speaking, the results of laboratory experiments may be less likely to be true outside of the lab (i.e., may have lower external validity) because the behavior of people in the artificial setting may differ from the actions they would have undertaken in the real world.[9] For example, Starmer (1999) indicated that the main line of criticism against the experimentalist economists is that the "results may be 'spurious' because of the 'artificial' context in which they are generated" (p. 4). According to List (2007, p. 34), "field experiments can alleviate typical criticisms of results from laboratory experiments by showing that such results have broader applicability than first believed in certain instances." A good example supporting the stance that lab experiments in economics lack generalizability are the exceptionally high estimates of discount rates (cf. Benhabib et al. 2010). While this fact can be assigned to the population of participants (usually students), there may be other systematic differences distorting drawing informative conclusions.

Steffen Andersen et al. (2010) directly compared the results of laboratory and field experiments. The evaluation of laboratory experiments in comparison to field experiments, relying on conducting the same lottery experiment, has been successful: the average risk aversions estimated on the basis of the two experiments is the same despite a higher variance of choices undertaken in the field. Andersen et al. (2010) explained the difference indicating different samples: while the lab experiment has been conducted on students, the participants of the field experiment has been randomly chosen from the Danish adult population. On the basis of these results, the economists concluded that "[t]he homogeneity of the university student population limits the ability of laboratory experiments to detect the preference heterogeneity that is present in the broader population" (Andersen et al. 2010, p. 222). List (2006), on the basis of comparison of market and laboratory behavior of experiment participants, concluded that in the controlled setting of laboratory experiments, people behave more pro-socially in comparison to their actions in the field driven by self-interest. In light of these results, field experiments seem to be more relevant for policy, but they are usually more expensive and time consuming. Also, their advantage is that they allow economists to construct any experimental situation and isolate factors under study that improves the estimation of average treatment effect (List 2007).

The methodological superiority of randomized field experiments has led to posing the question of whether field experiments differ significantly from randomized controlled trials (RCTs) as practiced in medical clinical trials. Judith Favereau (2016) indicated that one of the differences is (sometimes practiced) use of nonrandom assignments to groups. Furthermore, lab experiments in economics can rarely be double-blind since those who receive 'treatment' (e.g., educational or financial aid) are aware of it and therefore (i.e., due to knowledge of getting help rather than the help itself) may behave differently. However, Favereau (2016) indicated as the most significant difference the fact that field experiments can be only compared to the phase III of clinical trials in medicine (i.e., randomized controlled trials) while lacking the preclinical phase (in vitro and animal studies) and the phase II of clinical trial (tests on healthy humans aimed at establishing dosing). Based on this, Favereau established that the field experiments as practiced in economics could only be considered analogical to RCTs but not identical. David Teira and Julian Reiss (2013) extended the list of objections by adding the problem of self-selection bias: risk-averse individuals may be unwilling to participate in randomized experiments due to being afraid of getting into the control group. Given that risk-taking/risk aversion is one of the factors that can have a significant effect on economic success, self-selection bias may lead to a situation where risk-taking individuals take part in experiment, and therefore researchers observe better results. Based on this, Teira and Reiss (2013) claimed that RCTs but not field experiments can be considered as impartial method of causal inference.

I need to take issue with these claims. While I agree that some cases of field experiments in economics do not use double blinding[10] or even randomization, it is in principle possible to, for example, deliver a course/additional teaching to the treatment group and fake course to control group to introduce double-blind procedure. Similarly, random assignment is possible and should always be used.[11] Even if it is true that only risk-taking individuals participate in randomized field experiments, such studies can test the effectiveness of policy (intervention) within a group of risk-takers and, if properly designed and conducted, they offer internal validity. Furthermore, Favereau (2016) did not take into account that the first two phases of clinical trials[12] are used to gather background knowledge allowing for choosing the right treatment and dose, but their results are not used to evaluate effectiveness of treatment. This is exclusively the role of the phase III trials; i.e., the RCT (and, usually two such trials). As Karl Peace and Ding-Geng Chen (2010) put it, "[t]he primary objectives of the Phase III program are to confirm the effectiveness of the drug" (p 75). Since the first two phases are not conducted in the case of field experiments in economics, the background knowledge used to plan intervention needs to come from other types of evidence (e.g., anecdotal evidence or theoretical conjecture). However, this, on its own, does not undermine the use of randomized laboratory experiments as a tool for evaluating the effectiveness of microeconomic interventions. Therefore, in regards to evaluating policy, randomized field experiments could in principle be as strict as RCTs in medicine, at least if one considers ideal laboratory experiments and is concerned with internal validity.[13]

Another question which has recently received considerable attention in the methodological literature is related to the use of experimental evidence for policy. The debate addresses the question of when an experimental result can be extrapolated into a policy setting. This issue is known as the problem of extrapolation. The problem has been raised by Donald Campbell (1957), who coined the concepts of external and internal validity. The use of these notions differs across the philosophical literature. External validity is either taken as features of an experiment (i.e., it is an experiment that is valid) or an experimental or causal claim. For example, Francesco Guala (2003) took external validity as a feature of experiments rather than as causal claims describing experimental results. I believe that it is more fruitful to discuss the validity of claims rather than experiments, because one can produce many claims from one experiment and these claims can differ in the degree of their (both internal and external) validity. Hence, claiming that a causal claim is internally valid means that the evidence at hand is sufficient to conclude that the intervention of researchers conducting an experiment is responsible for producing the effect. On the contrary, the external validity of a causal claim denotes that it applies to or appropriately describes a causal structure of other settings. According to Guala (2003), experiments are externally valid if and only "if *A* causes *B* not only in *E*, but also in a set of other circumstances of interest, *F*, *G*, *H*, etc." (p. 1198).

However, the concepts of internal and external validity have been criticized for leading to poor evidential standards (e.g., Reiss 2019; cf. Jiménez-Buedo and Miller 2010). While the two notions may indeed cloud the issue at hand, Campbell (1957) has raised an important issue related to the use of experimental evidence: a causal claim that is true about what has happened in an experimental setting (i.e., is internally valid), may be false about a causal structure of a policy setting (i.e., externally invalid). This is the case especially in the domain of social sciences because the process of decision-making in humans may depend on self-conscious choice and therefore be less predictable than the behavior of other animals. Second, human behavior can be context dependent to such a degree that experimental results may not be replicable outside of artificial setting (Starmer 1999, p. 7).

Maria Jiménez-Buedo and Miller) summarized Campbell (1957) justification for putting forward the notion of external validity as follows: the "main purpose was . . . a matter of calling attention to the fact that randomization alone could not satisfactorily deal with all the possible interferences from extraneous factors that could potentially threaten the soundness of causal experimental conclusions." For the last 50 years, the problem of "generalizing results from laboratory to non-laboratory conditions" (Guala and Mittone 2005, p. 495) has been addressed by many philosophers and philosophically minded social researchers. Reiss (2019) offered a useful summary of the existing solutions to the problem of extrapolation present in the literature.

First, simple induction relies on taking experimental results as applicable to policy setting as long as convincing arguments contrary to this view are not present. Second, the generalization of experimental results can be considered an example of analogical reasoning based on analyzing the similarities between the experimental setting and the target population. Third, comparative process tracing is a method relying on the comparison of mechanisms between the experimental and target phenomena. Fourth, external validity can be ascertained by engineering the policy target; i.e., modifying relevant factors in order to obtain similarity to the experimental setting). Francesco Guala (2010, p. 1076) argued that comparative process tracing and analogical reasoning are, in fact, the same approach to the problem of extrapolation, which is based on comparing the similarity of causal structures between experimental and target settings. Finally, Reiss (2019) offered a new, 'contextualist' approach based on addressing the question of "what we need to know in order to make a reliable causal inference about a target system of interest" (p. 3115). According to Reiss, addressing policy questions should start from putting forth causal hypotheses and specifying what evidence is needed in support of that hypothesis.

Even though external validity is usually identified with features of a claim (being generalizable) or an experiment, I suppose that external validity could also be contextualized so that a claim is externally valid in regard to context C_1 but not generalizable to some other context C_1. Based on Reiss' (2019) solution, I differentiate between populational and extrapopulational extrapolation, and argue that experiments should be conducted while having in mind the population targeted by an intervention under consideration so that one does not extrapolate beyond the sampled population (cf. Section 6.3). Nevertheless, now I focus on discussing the methods of causal inference related to the manipulationist notion of causality.

6.2 Experimental and quasi-experimental research designs in economics

In this section, I analyze four case studies of research methods that allow for uncovering relations invariant under interventions. The number of economic studies that use experimental or quasi-experimental research design (design-based econometrics) is rapidly growing (Hamermesh 2013; Meyer 1995). The following two studies exemplify quasi-experimental research design. Doyle's (2007) study of the influence of foster care on children's wellbeing and income uses instrumental-variable design to construct quasi-experiment (Section 6.2.1). Pop-Eleches' (2006) analysis of the introduction of abortion ban in Romania employs regression-discontinuity design to study its influence on children's socioeconomic wellbeing. The latter two studies instantiate two experimental approaches. Hussam et al. (2008) conducted a laboratory financial market experiment to study how the experience of market participants influences the emergence of bubbles. Finally, I analyze the case of the gold standard of causal evidence. Dupas and Robinson (2013) conducted randomized field experiment to address the question of why saving propensity among the poor is low. In the following Section 6.3, I analyze what types of policymaking are justified by this evidence and deal with the problem of extrapolation.

6.2.1 Instrumental variable (IV) estimation as a quasi-experiment

Joseph Doyle (2007) employed the instrumental-variable estimation to assess the effects of foster care on children's wellbeing. On the basis of the study, Doyle (2007, p. 1583) concluded that "the results suggest that children on the margin of placement tend to have better outcomes when they remain at home, especially older children." In contrary to other research methods, economists using quasi-experimental research design do not restrain from explicit causal language: the discussion of causal effects (not associations or determinants as in the case of some econometric studies, cf. Chapters2–3) is present throughout the whole paper.

Before Doyle's study, the knowledge of the effects of foster care on children's wellbeing, health, and income had been limited. Direct comparisons of wellbeing of children experiencing foster care and those growing up in their families lead to misleading results, since the children taken from their homes (due to family problems) are at risk of experiencing socioeconomic difficulties in adulthood. Therefore, direct comparison of wellbeing of children growing up in their families and foster care is impossible. This is the case due to the problem of endogeneity and selection bias. Being placed in foster care is caused by children's characteristics; i.e., behavior and family background that also influence their wellbeing in adulthood. Furthermore, Doyle (2007) pointed out that no long-term data are accessible for the children who have experienced being investigated for either abuse or neglect but have remained with their families.

To estimate the effects of foster care on children, Doyle (2007) used the feature of the foster care system that provides quasi-randomization. Specifically, child protection investigators are assigned to children on rotational basis in a way that equalizes workload for each investigator. Furthermore, the decisions regarding taking children from family and assigning them to foster care depend on investigator's commonsense judgment rather than explicit rules (p. 1588), and therefore they differ considerably. Some investigators of child abuse cases tend to place children in foster care, while others decide to leave children with their biological parents. The quasi-random assignment of cases (children) to investigators allows for estimating the difference in outcomes (local average treatment effect, or LATE).

Given that LATE is the difference in outcomes for children who have received the treatment and would not have received the treatment even though they are eligible (e.g., abused sufficiently to justify sending to foster care), the definition of causality accepted implicitly by Doyle seem to be connected with a counterfactual formulation of manipulationist theory. To proceed, Doyle (2007) constructed an instrumental variable denoting the tendency of an investigator to place children in foster care. The tendency is defined as a ration of children placed in foster care divided onto the average ratio for all investigators. The IV is (statistically) independent from the characteristics of children that has been tested by a linear regression model. This allows for estimating marginal treatment effect (MTE). MTE denotes the benefit (or harm) from treatment for the individuals being precisely at the threshold of treatment (putting into the foster care). This, on the contrary, indicates that the study allows uncovering what action brings about the effect of improved children's wellbeing. As Doyle (2007, p. 1589) put it, "the results will consider the effect of assignment to different types of case managers, categorized by their rate of foster care placement, on long-term child outcomes."

What is vital for the policy implications of Doyle's (2007) quasi-experimental study is that the IV estimation allowed for estimating marginal treatment effect; i.e., the effect of being placed in foster care for children who are at the border between being left at their family home and being taken out. While the author conducted a robustness check that allows for concluding that this result is representative for a larger group of children (those who are not precisely at the limit), the effect may be different for those who are, according to contemporary standards, unlikely to be put in foster care. On the other hand, those unfortunate cases who strongly suffer in their familial homes may benefit from foster care (despite the estimated negative effect of the intervention) due to the severe implications of staying at home (e.g., homicide, serious physical abuse, etc.). Furthermore, the study used data of all children investigated for abuse in Illinois in chosen years and therefore allowed for estimating the effects of intervention (locating in foster care) that are representative for that population. Doyle (2007) observed some geographical heterogeneity in the sample (children in some counties are more likely to be put in foster care), suggesting that the population is not homogenous and this may undermine extrapolating this result to other states (cf. Section 6.3).

The relata of Doyle's causal claim are variables. Two types of variables occur. First, digital variables could be interpreted to represent either events (such as becoming teen mother) or features of instances (e.g., being Hispanic or white). Second, variables stand for measurable values (e.g., income). Overall, the variables denote attributes of instances or characteristics of children and investigators, and hence, the relata seem to be features of phenomena.

The trust in the causal claim to be implementation neutral depends on whether the instrument allows for 'as-if' random allocation between control and treatment group. As Joshua Angrist et al. (1996, p. 454) put it, "the strong assumptions [are] required for a causal interpretation of IV estimates". If these assumptions are not fulfilled, then the usual criticism of econometric models applies (cf. Section 3.3). However, given the successful quasi-randomization, Doyle's (2007) evidence justifies the belief that the relation is invariant under intervention.

6.2.2 Natural experiments: regression discontinuity design

While the IV estimation, exemplified previously, is a popular method of analyzing cross-sectional data, time series analysis has been dominated by regression-discontinuity design (RDD). This approach is based on the comparison of data from periods just before and just after an intervention. The difference between the two groups mimicking control and treatment groups is considered as resulting from the intervention under consideration. A slightly different approach would be to measure differences in trend before and after interventions, what is known as difference-in-differences design. Even though these quasi-experimental studies are representative for design studies in macroeconomics, I have chosen the paper of Cristian Pop-Eleches (2006) focusing on the influence of an abortion ban on children's educational attainment and labor market outcomes because this study is an excellent example of how quasi-experimental research design can suffer from confounding.

Pop-Eleches (2006) studied the influence of abortion ban introduced in Romania in 1966 on children's economic success and educational attrition. His main conclusion is that "children born after the ban on abortions had worse educational and labor market achievements as adults" (p. 744). However, a straightforward application of the regression discontinuity design, based on the comparison of pre-intervention and post-intervention samples (children born in the period of a few months before and after the ban was introduced) leads to the opposite conclusion. Such a simple analysis relies on estimating the following linear regression (Pop-Eleches 2006, p. 753):

$$OUTCOME_i = \alpha_0 + \alpha_1 \star after_i + \varepsilon_i$$

Where:

$OUTCOME_i$ = digital variables measuring educational attrition and labor market outcomes for adults
α_0 = the probability of succeeding for children not affected by abortion ban
α_1 = the influence of being born after the ban on the probability of succeeding for children
$after_i$ = *i*-th children affected by the abortion ban (digital variable)
ε_i = the error term

Surprisingly, such a simple design of the natural experiment led to obtaining results opposite of the expected outcomes of the abortion ban. The estimated values ($\alpha_0 > \alpha_1$) showed that, on average, children after the abortion ban were more likely to have better education and higher position on the labor market. This result opposes previous studies of the influence of liberalization of abortion law in United States and other countries that have reported the improvement of children's situation after the increase of the number of abortions (e.g., Levine et al. 1996; Koch et al. 2012). However, adding additional explanatory variables to the regression (i.e., observable characteristics of children such as familial and economic background) leads to the reversal of the preliminary finding.

$$OUTCOME_i = \beta_0 + \beta_1 \star after_i + \beta_2 \star X_i + \varepsilon_i$$

Where:

X_i = vector of variables characterizing *i*-th child's background
β_0 = constant
β_1 = the influence of being born after the ban on the probability of succeeding for children
β_2 = the influence of child characteristics on the probability of success

When the regression controlling for confounders is estimated, the abortion ban has *negative* and not positive effect; i.e., $\beta < \beta_1$. Pop-Eleches (2006) explains the reversal of the average treatment effect by the fact that educated women are the group most affected by the ban (because abortions before 1966 had been most frequent among the members of this group). Given that children born by higher-educated women have a higher chance of finishing their education, and unequal influence of abortion ban on different social groups, the result of the simple regression (not controlling for confounding) delivers results contrary to the actual influence of the abortion ban on socioeconomic factors.

The results of Pop-Eleches (2006) show how the use of quasi-experimental research design (RDD, in this case) may result in spurious results in cases when there are confounders that bias the estimate and create non-random assignments to treatment and control groups. What follows, ascribing an observed effect of an intervention to it, requires careful examination of whether any confounding factors can be present. If either there are no confounders or their effects are taken into account, then regression-discontinuity design allows for uncovering relation that is invariant under intervention. Taking into account that the abortion ban is a human action and the effects follow from that actions, a version of Menzies and Price's (1993) agency theory of causality generalized into a probabilistic context seems to be a good candidate for the views *implicitly* accepted by the economists using RDD.

However, the design only allows for uncovering average treatment effect. Similar to the previous example, predicting the effects of the intervention for each potentially affected child is impossible. This gives a hint that economists using quasi-experimental research design accept a version of manipulationist definition focusing on type-level relation.

Furthermore, the analysis of the effects of the abortion ban indicates that even evidence supporting manipulationist claims cannot be extrapolated into other contexts without caution because other causal factors can play a role. In the case of abortion legislation, the opposite effects can be observed for the American and Romanian populations, as long as one does not control for confounders. The liberalization of abortion law in the United States reduced crime rates and had a positive impact on the socioeconomic outcomes of affected children because the procedure is mostly used by mothers who have social and economic difficulties themselves. On the contrary, the Romanian ban on abortion affected mostly mothers living in cities.

Introducing statistical control of confounders requires knowledge of the factors that influence effects of intervention. Given that our knowledge of other important factors may be limited (some confounders stay unknown) or fallacious (controlling for 'confounders' that are not causally related, but only correlated to outcomes, may lead to spurious causal claims), quasi-experimental research designs deliver results less reliable than the gold standard of causal inference. This is the case, because randomization allows, at least in principle, for construction of control and treatment groups influenced by confounders in the same way so that the difference can only result from the intervention.

6.2.3 Laboratory experiments

According to orthodox neoclassical economics, prices on financial markets reflect the fundamental value of assets. While this view has faced considerable criticism, especially after the 2007–2008 financial crisis (e.g., Krugman 2009), it persists despite the presence of contradicting empirical evidence. One of the phenomena in disagreement with market efficiency is bubbles. The volatility of stock prices cannot be explained by changes in real values such as predicted stock returns (West 1988; Vogel and Werner 2015), and this evidence indicates that bubbles do happen. Unfortunately, high uncertainty of financial predictions and epistemic inaccessibility of the decision-making process of market participants make it impossible to definitively describe a price trend as a bubble before it bursts. What follows, studying the development of bubbling markets in the world, has serious epistemic limitations.

For this reason, economists construct asset markets in a laboratory and hope that changing the conditions under which market participants undertake their decisions and observing effects will shed light on how these markets work and why they diverge from the ideal of efficient market. The study of Reshmaan Hussam et al. (2008) is a representative example of this type of laboratory experiments. Game-theoretic experiments are another common type of laboratory experiments (Maziarz 2018). In case of this type of laboratory experiments, economists construct game settings (e.g., the prisoner's dilemma or the ultimatum game) to test predictions of the rational expectations model and observe how actual decisions diverge from this ideal.

Hussam et al. (2008) conducted a series of laboratory market experiments aimed at assessing how the learning and experience of market participants influence price bubbles. Similar to the quasi-experimental studies considered previously, the analysis entails explicitly causal talk throughout the paper. The two causal claims formulated by the authors state that "[e]xperience reduces the amplitude of a bubble significantly" and "[e]xperience significantly reduced the duration of a bubble" (Hussam et al. 2008, pp. 933–934).

To obtain these results, the economists conducted the canonical asset market experiments on a group of undergraduate students. This involved an asset that lasts for 15 trading sessions and pays a random dividend at each session (the distribution of dividends is constant throughout all 15 sessions). Hence, the fundamental value of the asset should start at the value equal to 15 * *expected value of dividend* and deteriorate linearly. At the start of experiments, participants in a given round received one of three portfolios including cash and the asset in different proportions. In order to manipulate the level of experience of the participants, Hussam et al. (2008) divided participating students into two groups. The first group took part in two 15-session rounds. These were learning sessions aimed at obtaining experienced students, and their results were not taken into account.

Later, these experienced students were randomly divided so that they participated in trading sessions with either inexperienced students or students after one or two rounds. Furthermore, Hussam et al. (2008) modified the variance of dividend and the cash value owned by participants (rekindle treatment). These two interventions have been recognized in other experiments to promote bubbles (e.g., Lei et al. 2001). The purpose of these modifications was to test for the robustness of the results. To assess the degree to which asset markets in each round bubbled, Hussam et al. (2008, p. 932) developed a few measures of bubbling such as amplitude of prices and duration of periods when the asset was traded at price levels different from its fundamental value. The results indicated that the experience of experimentees has a negative effect on bubbles, reducing their amplitude and duration. The effect of experience and the rekindle treatment was similar in size. This sheds light on the fact that bubbles can appear despite the experience of traders because of changes in market environment such as dividend payoffs.

What does it mean that the experience reduces bubbles on the asset market? What is the definition of causality presupposed by Hussam et al. (2008)? Reiss (2019, pp. 3113–3114, emphasis in original) argued that "[a]n ideal randomised trial implements an ideal intervention" but "no *real* randomised experiment implements an *ideal* intervention." The reasons, according to Reiss, are unequally distributed confounders among treatment and control groups, failure at blinding, and different dropout rates between treatment and control groups. Donald Gillies (2018, Appendix 1) extended the list by arguing that treatment interventions affect outcomes by other paths than by the variable directly targeted (e.g., by the placebo effect that influences measured outcome; i.e., disease but not by the mechanism targeted by a drug under test). The interventions in laboratory experiments also do not fulfill Woodward's highly technical definition of intervention. Let me exemplify this view with the case of Hussam et al.'s (2008) modifications of the experience level of experimentees. Specifically, Woodward (2005 [2003], p. 98) requires the intervention *I* to act "as a switch for all the other variables that cause *C*." In other words, intervention on the level of experience should screen off all other factors shaping the asset market experience of students. This condition is definitely not fulfilled. Obviously, students may have gained some experience from participating in Finance 101, investing their savings in financial markets, or merely reading about canonical asset market experiments that are a standard setting for studying market (in)efficiency in laboratory settings. Therefore, Hussam et al.'s (2008) action shaping the level of experience of student participants is not 'intervention' in the Woodwardian sense.

However, what the laboratory market experiment allows for discovering is the relation between the treatments (modifying the level of experience and the rekindle treatment) and outcomes of interest. In other words, the laboratory asset market allowed to infer that raising the level of experience of market participants reduces the magnitude and duration of market bubbles. Given this and considering the treatments in laboratory experiments in economics (and, in fact, all other experimental sciences) lie within the scope of human capabilities, then Menzies and Price's (1993) agency theory delivers a good candidate for the definition presupposed by experimenters. The relata are variables standing for features of phenomena. Under the action of the experimenters, some features (level of experience of market participants, dividend distribution) are modified, and these changes lead to changes in the propensity of the market to bubble. These actions are meant to bring about more efficient market pricing of the asset.

Unfortunately, causal claims are true only within the laboratory. The epistemic situation of these claims can be compared to theoretical models. In Chapter 5, I argued that mechanistic knowledge of one mechanism is insufficient to predict the effects of interventions. Given that theoretical models of economic mechanisms isolate (and idealize) single mechanisms, and that there are many mechanisms operating in the world at the same space and time, predictions based on a model of one mechanism, despite being true within the model world, will prove false in the actual reality due to the presence of external influences. This view, if correct, makes testing the accuracy of economic models of mechanisms by comparison of predictions to econometric results impossible (cf. Maziarz 2019). However, the example of Hussam et al.'s (2008) laboratory experiment shows that the verification of the accuracy of mechanistic models can proceed by constructing artificial settings in the laboratory. In this way, economists can construct artificial laboratory market as it is described by the assumptions of a theoretical model and test whether the decisions undertaken by economic agents are in agreement with the predictions deduced from the model under consideration. The similarity of laboratory experiments and theoretical models also implies that the 'experimental closure' (the isolation of one mechanism from external factors) have severe and detrimental effects on the use of laboratory results for policymaking (cf. Section 6.3).

6.2.4 Randomized field experiments

Randomized field experiments are not susceptible to this criticism because they test whether a treatment brings about an expected outcome in the field and not under the experimental closure. While the topics studied by means of randomized field experiments range from the topics strictly located within the scope of economics such as the effects of basic income on the behavior of recipients on labor markets to the intersection of economics and medicine (e.g., the effects of anti-mosquito bed nets), the majority of studies using this design aim at testing the effectiveness of some policies improving economic growth. Experiments are usually conducted in developing countries. This practice allows for reducing costs of experimentation, but also limits the chance of extrapolating the results (cf. Section 6.3). The field experiment of Pascaline Dupas and Jonathan Robinson (2013) is a representative example of this type of research.

The trial of Dupas and Robinson (2013, p. 163) concludes that "simple informal savings technologies can substantially increase investment in preventive health and reduce vulnerability to health shocks." Contrary to clinical trials in medicine, which randomly allocate participants among treatment and control groups, Dupas and Robinson randomized local savings organizations (rotating savings and credit associations, or ROSCA) in one of the Kenyan administrative regions. These organizations have been divided into five groups, four of which received different treatments and one became a control group.[14] The four treatment groups received the following treatments. Participants belonging to the first group received a metal box (a piggy bank) delivering space to keep money at home in a considerably safe place. The second group received the same piggy bank box, but were refused the key to it so that the money could be accessed only after exceeding their saving goal. The third group saved money to a 'health pot' managed by ROSCA that bought health products on behalf of (randomly chosen) recipients of the funding. The fourth group got access to a health savings account. All the treatments proved effective and raised the ratio of people having health savings from over 60% to over 90% in six- and 12-month periods.

Dupas and Robinson (2013) had also considered treatment 'technology' and related them to features such as access to storage or social commitment. These allowed them to study econometrically the influence of participants' characteristics with the view to relate the effectiveness of different treatments to these characteristics. Given that this aspect of Dupas and Robinson's (2013) study heavily overlaps with observational research, I exclude it from the analysis. Therefore, the main conclusion of the experimental design is that delivering tools such as saving boxes or savings accounts raises poor peoples' propensity to save.

This conclusion is a type-level causal claim. It allows inferring that introducing one of the treatments will raise saving propensity of targeted population (under the assumption that the population sufficiently resembles the people living in Kenyan rural areas), but it does not allow the inference that each recipient of the treatment will benefit from it. As in the previous cases of experimental research, the relata of the causal claim are variables standing for features of phenomena. The treatment used by Dupas and Robinson (2013) is in disagreement with Woodward's (2005[2003], p. 98) third axiom describing the ideal intervention *I*. According to Woodward, interventions should influence the effect *E* by no other path than *C*. The discussion of the role of health savings, applied to both control and treatment groups, may, at least in principle, affect *E*; i.e., the health savings propensity of the poor without changing *C*. Given this, the experimental design of Dupas and Robinson (2013) does not allow for uncovering the results of the interventions in the Woodwordian sense, but it nevertheless allows for discovering that delivering saving tools is an effective means to raise the health savings level among the poor in rural Kenya, and therefore Menzies and Price's (1993) formulation of agency theory can be indicated as a concept of causality implicitly presupposed by the authors. Ahead, I address the questions of whether the relation invariant under intervention can be applied to other settings stays open.

6.3 Is extrapolation from experimental studies always problematic?

As I have argued, both actual experiments (laboratory and field experiments) and quasi-experimental research designs (if correctly applied) allow for finding relations that are invariant under interventions; i.e., deserve the causal label even if one accepts a version of the manipulationist view on causality. The invariance under interventions of uncovered relations, in principle, allows for conducting interventions that change the relata of causal claims (causes). Given the invariance, these changes lead to (at least on average[15] or under the *ceteris paribus* clause) the effects being in agreement with causal claims.

However, the problem of extrapolation allegedly occurs: while causal claims are true for the sample population under experimental (or quasi-experimental) study, their accuracy in the target population stays unknown. Many different strategies of dealing with the problem of extrapolation have been offered, but their success seem to be limited so far since they usually require knowing the target that suffices for establishing a causal claim on its own. This issue is known as the extrapolator's circle (cf. Section 6.1.2). I want to offer a solution to the problem of extrapolation based on differentiating between two types of extrapolation that seem to be entangled in the philosophical debates. This solution is in line with Julian Reiss' (2019) advice to use contextualized evidence for policy, but develops his approach by offering more practical guidance.

6.3.1 Causal structures and populations

First, let me explain why the problem of extrapolation is a hot topic in the philosophy of social sciences (e.g., Steel 2004), but not in physics or medicine. In the natural sciences, the reality is assumed to be unified in the sense that experiments conducted in one place are likely to give results similar to the same studies conducted elsewhere. For example, measuring the time of falling apples in Finland and Poland to calculate the approximation of earth's gravitational force leads to obtaining similar results. In a similar vein, clinical trials aimed at testing for the efficiency of a drug in humans are believed (sometimes wrongly, cf. Stegenga 2018) to lead to similar conclusions. This is the case because it is believed that both gravitational force in Finland and Poland, and the mechanisms at work in Finnish and Polish patients, are sufficiently similar.

On the contrary, the realm of social sciences is considered to be more fragmented regarding populations. Therefore, while the evidence that a drug cures Finns is accepted as justifying the claim that the drug will be efficient in curing Poles, using experimental evidence from Finland about the domain of social sciences in Poland raises the problem of extrapolation. This is the case even though the results of clinical trials also need to be extrapolated into the target population because studies are conducted on a sample of the population. However, having clinical evidence that vitamin C cures the common cold, no one asks whether this result applies to (can be extrapolated into) the populations of economists or philosophers of science.

6.3.2 The two types of extrapolation

This is the case because this type of extrapolation (used in the natural sciences) does not cross the borders of populations with, possibly, different mechanisms at work. On the contrary, given that statisticians have dealt with this problem (of sampling) occurring in natural sciences since the beginning of the eighteenth century (Stigler 1986), we are able to quite well assess the likelihood that a result obtained on a sample can be extrapolated to the population. Let me call this type 'populational extrapolation,' which denotes extrapolating from a representative (random) sample of a population into that population. An intervention justified by evidence from the populational extrapolation can fail for the following two reasons. First, the sample may be biased. Even though using a random sample in principle warrants its representativeness (in the limit), neither real-life samples nor populations are infinite, and therefore sampling may lead to obtaining biased samples – and hence results not representative for the whole population. Second, even this (safer) version of extrapolation requires extrapolating in time. If causal structure of that population changes in time, then interventions may lead to unexpected outcomes.

The other type of extrapolation ('extra-populational' extrapolation) is much riskier. Extra-populational extrapolation is a process of taking experimental evidence obtained from a sample of one population and interpolating results into some other population. In the natural sciences, extra-populational extrapolation would occur if one tested the efficiency of vitamin C in curing the common cold in a non-human animal model and extrapolated results onto the population of economists and philosophers of science. I read the literature on the problem of extrapolation and external validity as being primarily concerned with the extra-populational extrapolation. To warrant the eligibility of evidence extrapolated extra-populationally, one needs to ascertain that the causal structure in both populations is sufficiently similar so that the same interventions produce the same outcomes. Given that the knowledge of causal structure allows for predicting the outcomes of interventions, the extrapolator's circle occurs. Considering the epistemic risks involved in extra-populational extrapolation, the US Food and Drug Administration requires testing drugs in humans (by means of RCTs) in order to avoid these risks (Ciociola et al. 2014).

6.3.3 Toward contextualized experimental evidence

My solution to the problem of extrapolation in the social sciences is similar. Policymakers should avoid extra-populational extrapolation and use evidence from the target population instead. While such a solution may seem to be very expensive, since it requires gathering experimental evidence for every intervention and each population, the costs of failed interventions may incur even more severe losses.[16] The main problem with the distinction between populational and extra-populational extrapolation in the domain of social sciences is that strict borders between populations may be nonexistent. Contrary to the natural sciences, where differentiating between humans and rats suffering from common cold may be easy, populations in the social sciences may be difficult to define. For example, we do not know whether children at the risk of foster care in New York differ significantly from such children in Chicago, and therefore deciding whether using Doyle's result to modify the decisions of social workers in New York is an example of populational or extra-populational extrapolation may be ambiguous.

However, defining the borders of populations in the policymaking practice may actually be more comfortable than it seems, since population under interest can easily be defined as all people influenced by a policy under consideration. Using as evidence for policymaking in population *A* an experiment on a random sample of population *A* requires only populational extrapolation. Therefore, policymakers should first conduct a 'pilot policy' or 'pilot intervention' – a randomized field experiment or a quasi-experimental study of a representative sample of population (random sample) – and take the results as an indicator of whether the intervention under study will be efficient in the target (i.e., populationally extrapolate the result). Such an approach resembles the way pharmaceutical companies conduct clinical trials (RCTs) and allows for refraining from extra-populational extrapolation, which raises the problem of extrapolator's circle.

My solution to the problem of extrapolation can easily be applied to evidence from laboratory field experiments. The possibility of using quasi-experimental designs to address policy-specific questions depends on the accessibility of data. However, the use of laboratory experiments as evidence for policy, due to the artificial setting, seems to always involve extra-populational extrapolation. Given this, and taking into account the epistemic risks connected to this type of extrapolation, laboratory experiments may be of limited use as direct evidence for policymaking. However, it does not mean that economists should refrain from this type of study. Laboratory experiments play an important role in economics, since they allow for testing assumptions and predictions of theoretical models. This is the case because they allow for observing how a single causal factor shapes decision-making of economic agents (cf. Chapter 5). Furthermore, laboratory experiments, being relatively cheap, can deliver guidance for planning policy-relevant field studies.

While the evidence justifying conclusions in agreement with the manipulationist notion of causality is the gold standard, it neither justifies any causal claims nor warrants all interventions to be effective. In particular, all studies considered previously estimate the average effects of interventions for some group. Therefore, the causal claims are invariant under intervention only at the aggregate level. Manipulationist evidence, as delivered by economics design-based studies, at the level of single instances (token-level causality), is only capable of delivering probabilistic manipulationist claims of how an intervention under consideration changes the likelihood of the effect. Therefore, translating type-level manipulationist causal claims into token-level counterfactuals (describing the effects of interventions at the level of instances) is not justified.

6.4 Manipulationist evidence and interventions

In this chapter, I have argued that Woodward's interventionist theory uses too strict anotion of intervention to be applicable in research practice. In particular, the condition of screening off and only one causal path from intervention to effect cannot be met in the research practice of economics. Therefore, Woodward's definition of causality can serve as an ontological view. In the research practice, economists can assess the effects of 'interventions' understood as human actions. Therefore, the agency theory of Menzies and Price (1993) in its probabilistic formulations (i.e., stating that actions raise the probability that effects follow) seems to be presupposed by economists using experimental or quasi-experimental research designs.

However, inferring relations invariant under interventions does not imply that they can be employed in all types of policymaking. Causal claims based on the experimental and quasi-experimental research designs hold only at the level of populations. At the singular level, interventions only modify the probability that an effect will follow, but do not allow for concluding that a causal claim is deterministically true for each case. Therefore, translating manipulationist causal claims into the counterfactual notion may lead to unexpected policy outcomes.

Furthermore, using experimental results for policy in the social sciences is criticized for being susceptible to extrapolator's circle. I have argued that the philosophical debates on the problem of extrapolation entangle two notions of extrapolation: populational and extra-populational extrapolation. While the former type is unproblematic and relies on using the statistical theory of sampling for estimating the likelihood that an experimental result is representative, the success of the latter type of extrapolation depends on the (unknown) level of similarity of causal structures of experimental population and the target.

Notes

1 In Housman and Woodward's (1999, 2004) dictionary, the 'wiggling' denotes an intervention on a variable (e.g., X)aimed at changing the variables following them on a causal path in acyclic graph (its effects; e.g., Y).

2 Considering the limited (not to say nonexistent) use of Bayesian networks in economics, I will limit the discussion to a brief summary. This is useful to understand the inspiration for Woodward's interventionist theory.

3 I need to notice that even if a system under study were linear and deterministic, the invariance of causal relations could be difficult to observe, given that other factors change simultaneously and influence the effect of intervention under consideration.

4 In *Making Things Happen: A Theory of Causal Explanation*, Woodward (2005 [2003]) distinguished between causal and non-causal (correlational) interpretation of structural equations, cf. Chapter 2.

5 Henschen specifies further conditions (2018, pp. 16–17). The detailed discussion of his definition and the argument that it is only partially adequate can be found elsewhere (Maziarz and Mróz 2019).

6 The meaning of mechanism presupposed by Judea Pearl differs from how the concept is defined within the mechanistic philosophy (cf. Chapter 5) and resembles the views of econometricians practicing the structural equation approach (Reiss 2013).

7 Randomized controlled trials deserve the label of 'gold standard' for they allow, at least if samples were infinite, for randomizing all confounding factors so that the difference between treatment and control group can be assigned exclusively to intervention. I discuss randomized field experiments and laboratory experiments later in the chapter.

8 The readers unacquainted with the basic methodological issues of experimental studies may be willing to skip to the following session.

9 The distinction between internal and external validity is sometimes labeled with the notions of efficacy (denoting effects within the trial) and efficiency (denoting effects generalizable into a policy target) (cf. Cartwright 2009).

10 The procedure that forbids both patients and clinicians to know whether they are/treat a member of the treatment/control group.

11 While I am aware that in some cases the random assignment to groups is corrected in order to get samples having the same characteristics as global population, I believe that this can lead to false results since correcting the distribution of known confounders may influence the distribution of unknown (and possibly unobservable) confounders.

12 Clinical trials usually consist of the following three phases. First, a new drug is tested on a small group of healthy people in order to check if it is healthy. Second, it is given to a group of patients suffering from a disease to verify if it can affect their condition. Finally, RCT is used to assess the effectiveness of the new drug.

13 The difference for *the use of results for policy* between RCTs and randomized field experiments come from different type of extrapolation, and I will elaborate on this issue in Section 6.3.

14 In order to test the effectiveness of different savings tools rather than what they all have in common (i.e., advising health savings), the control group has also received a (partial) treatment of a meeting with a member of the experimentation team that discussed the benefits of having health savings.

15 That is, the intervention raises the chances for the effect to occur.

16 Certainly, if a policy under consideration is of limited significance, then using extra-populational extrapolation may be justified on the basis of cost-benefit analysis.

References

Achen, C. H. (1986). *The Statistical Analysis of Quasi-Experiments*. Berkley: University of California Press.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior & Organization*, *73*(2), 209–224. DOI: 0.1016/j.jebo.2009.09.006

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455. DOI: 10.1080/01621459.1996.10476902

Angrist, J. D., & Krueger, A. B. (1992). *Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery* (No. w4067). National Bureau of Economic Research. DOI: 10.3386/w4067

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, *15*(4), 69–85.

Angrist, J. D., & Kuersteiner, G. M. (2011). Causal effects of monetary shocks: Semiparametric conditional independence tests with a multinomial propensity score. *Review of Economics and Statistics*, *93*(3), 725–747. DOI: 10.1257/jep.15.4.69

Angrist, J. D., & Pischke, J. S. (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. DOI: 10.1257/jep.24.2.3

Banks, J., & Mazzonna, F. (2012). The effect of education on old age cognitive abilities: Evidence from a regression discontinuity design. *The Economic Journal*, *122*(560), 418–448. DOI: 10.1111/j.1468-0297.2012.02499.x

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, *69*(2), 205–223. DOI: 10.1016/j.geb.2009.11.003

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*(4), 297. DOI: 10.1037/h0040950

Cartwright, N. D. (2001). What is wrong with Bayes nets? *The Monist*, *84*(2), 242–264.

Cartwright, N. D. (2002). Against modularity, the causal Markov condition, and any link between the two: Comments on Hausman and Woodward. *The British Journal for the Philosophy of Science*, *53*(3), 411–453. DOI: 10.1093/bjps/53.3.411

Cartwright, N. D. (2003). Two theorems on invariance and causality. *Philosophy of Science*, *70*(1), 203–224. DOI: 10.1086/367876

Cartwright, N. D. (2006). From metaphysics to method: Comments on manipulability and the causal Markov condition. *The British Journal for the Philosophy of Science*, *57*(1), 197–218.

Cartwright, N. D. (2007a). Are RCTs the gold standard? *BioSocieties*, *2*(1), 11–20.

Cartwright, N. D. (2007b). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511618758

Cartwright, N. D. (2009). What is this thing called 'efficacy'? In: Mantzavinos, C. (ed.) *Philosophy of the Social Science: Philosophical Theory and Scientific Practice* (pp. 185–206). Cambridge: Cambridge University Press.

Caughey, D., & Sekhon, J. S. (2011). Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008. *Political Analysis*, *19*(4), 385–408. DOI: 10.1093/pan/mpr032

Ciociola, A. A., Cohen, L. B., Kulkarni, P., Kefalas, C., Buchman, A., Burke, C., . . . & Fass, R. (2014). How drugs are developed and approved by the FDA: Current process and future directions. *The American Journal of Gastroenterology*, *109*(5), 620. DOI: 10.1038/ajg.2013.407

Collingwood, R. G. (2001 [1940]). *An Essay on Metaphysics*. Oxford: Clarendon Press.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Cooley, Th., LeRoy, S., & Raymon, N. (1984). Econometric policy evaluation: Note. *American Economic Review*, *74*, 467–470.

Dawid, A. P. (2010). Beware of the DAG! *Proceedings of Workshop on Causality: Objectives and Assessment at NIPS 2008*, San Diego: PMLR, 6, pp. 59–86.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, *48*(2), 424–455. DOI: 10.1257/jel.48.2.424

Doyle Jr, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review*, *97*(5), 1583–1610. DOI: 10.1257/aer.97.5.1583

Dunning, Th. (2012). *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139084444

Dupas, P., & Robinson, J. (2013). Why don't the poor save more? Evidence from health savings experiments. *American Economic Review*, *103*(4), 1138–1171. DOI: 10.1257/aer.103.4.1138

Engle, R. F., & Hendry, D. F. (1993). Testing superexogeneity and invariance in regression models. *Journal of Econometrics*, *56*(1–2), 119–139. DOI: 10.1016/0304-4076(93)90103-C

Engle, R. F., Hendry, D. F., & Richard, J. F. (1983). Exogeneity. *Econometrica: Journal of the Econometric Society*, 277–304.

Favereau, J. (2016). On the analogy between field experiments in economics and clinical trials in medicine. *Journal of Economic Methodology*, *23*(2), 203–222. DOI: 10.1080/1350178X.2016.1157202

Freedman, D. A. (2007). Statistical models for causation. In: Outhwaite, W. & Tumer, S. (eds.) *The Sage Handbook of Social Science Methodology* (pp. 127–146). London: SAGE Publications.

Friedman, M. (1994 [1953]). The methodology of positive economics. In: Hausman, D. (ed.) *The Philosophy of Economics: An Anthology*. New York: Cambridge University Press.

Gillies, D. (2018). *Causality, Probability, and Medicine*. London: Routledge.

Guala, F. (2003). Experimental localism and external validity. *Philosophy of Science*, *70*(5), 1195–1205. DOI: 10.1086/377400

Guala, F. (2005). *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press.

Guala, F. (2010). Extrapolation, analogy, and comparative process tracing. *Philosophy of Science*, *77*(5), 1070–1082. DOI: 10.1086/656541

Guala, F., & Mittone, L. (2005). Experiments in economics: External validity and the robustness of phenomena. *Journal of Economic Methodology*, *12*(4), 495–515. DOI: 10.1080/13501780500342906

Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, *51*(1), 162–172. DOI: 10.1257/jel.51.1.162

Hausman, D. M. (1983). Are there causal relations among dependent variables? *Philosophy of Science*, *50*(1), 58–81. DOI: 10.1086/289090

Hausman, D. M. (1998). *Causal Asymmetries*. Cambridge: Cambridge University Press.

Hausman, D. M., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *The British Journal for the Philosophy of Science*, *50*(4), 521–583. DOI: 10.1093/bjps/50.4.521

Hausman, D. M., & Woodward, J. (2004). Manipulation and the causal Markov condition. *Philosophy of Science*, *71*(5), 846–856. DOI: 10.1086/425235

Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century retrospective. *The Quarterly Journal of Economics*, *115*(1), 45–97. DOI: 10.1162/003355300554674

Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, *76*(1), 1–27. DOI: 10.1111/j.1751-5823.2007.00024.x

Hendry, D. (2000). Econometrics: Alchemy or Science? Essays in Econometric Methodology. Oxford: Oxford University Press.

Henschen, T. (2018). What is macroeconomic causality? *Journal of Economic Methodology*, *25*(1), 1–20. DOI: 10.1080/1350178X.2017.1407435

Henschen, T. (2020). Response to 'Response to Henschen: Causal pluralism in macroeconomics *Journal of Economic Methodology*, First View. DOI: 10.1080/1350178X.2020.1730575

Hogan, J. W., & Lancaster, T. (2004). Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Statistical Methods in Medical Research*, *13*(1), 17–48. DOI: 10.1191/0962280204sm351ra

Hoover, K. D. (1990). The logic of causal inference: Econometrics and the conditional analysis of causation. *Economics & Philosophy*, *6*(2), 207–234. DOI: 10.1017/S026626710000122X

Hoover, K. D. (2001). *Causality in Macroeconomics*. Cambridge: Cambridge University Press.

Hoover, K. D. (2008). Causality in economics and econometrics. *The New Palgrave Dictionary of Economics* (Vol. 1–8, pp. 719–728). London: Palgrave Macmillan

Hoover, K. D. (2011). Counterfactuals and causal structure. In: Illari, Ph. et al. (eds.) *Causality in the Sciences* (pp. 338–360). Oxford: Oxford University Press.

Hoover, K. D. (2012). Economic theory and causal inference. In: Mäki, U. (ed.) *Philosophy of Economics* (pp. 89–114). Amsterdam: Elsevier.

Hussam, R. N., Porter, D., & Smith, V. L. (2008). Thar she blows: Can bubbles be rekindled with experienced subjects? *American Economic Review*, *98*(3), 924–937. DOI: 10.1257/aer.98.3.924

Jiménez-Buedo, M. (2011). Conceptual tools for assessing experiments: Some well-entrenched confusions regarding the internal/external validity distinction. *Journal of Economic Methodology*, *18*(3), 271–282.

Jimenez-Buedo, M., & Miller, L. M. (2010). Why a trade-off? The relationship between the external and internal validity of experiments. *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia*, *25*(3), 301–321. DOI: 10.1080/1350178X.2011.611027

Koch, E., Thorp, J., Bravo, M., Gatica, S., Romero, C. X., Aguilera, H., & Ahlers, I. (2012). Women's education level, maternal health facilities, abortion legislation and maternal deaths: A natural experiment in Chile from 1957 to 2007. *PLoS One*, 7(5), e36613. DOI: 10.1371/journal.pone.0036613

Koopmans, T. C. (1950). When is an equation system complete for statistical purposes? In: Koopmans, T. C. (ed.) *Statistical Inference in Dynamic Economic Models*. New York: John Wiley and Sons.

Krugman, P. (2009). How did economists get it so wrong? *New York Times*. Retrieved from: www.nytimes.com/2009/09/06/magazine/06Economic-t.html?_r=1&pagewanted=print. Access: 18th March 2019.

Lalive, R. (2007). Unemployment benefits, unemployment duration, and post-unemployment jobs: A regression discontinuity approach. *American Economic Review*, *97*(2), 108–112. DOI: 10.1257/aer.97.2.108

Leamer, E. E. (1983). Let's take the con out of econometrics. *Modelling Economic Series*, *73*, 31–43.

Lee, D. S., & Lemieux, T. (2009). *Regression Discontinuity Designs in Economics*. Cambridge: National Bureau of Economic Research.

Lei, V., Noussair, C. N., & Plott, C. R. (2001). Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica*, *69*(4), 831–859. DOI: 10.1111/1468-0262.00222

LeRoy, S. F. (2016). Implementation-neutral causation. *Economics & Philosophy*, *32*(1), 121–142. DOI: 10.1017/S0266267115000280

LeRoy, S. F. (2018). Implementation neutrality and treatment evaluation. *Economics & Philosophy*, *34*(1), 45–52. DOI: 10.1017/S0266267117000219

Levine, P. B., Trainor, A. B., & Zimmerman, D. J. (1996). The effect of medicaid abortion funding restrictions on abortions, pregnancies and births. *Journal of Health Economics*, *15*(5), 555–578. DOI: 10.1016/S0167-6296(96)00495-X

Lewis, D. (1973). Causation. *The Journal of Philosophy*, *70*(17), 556–567. DOI: 10.2307/2025310

List, J. A. (2006). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of Political Economy*, *114*(1), 1–37. DOI: 10.1086/498587

List, J. A. (2007). Field experiments: A bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis & Policy*, *5*(2). DOI: 10.1086/498587

Maziarz, M. (2018). Causal inferences in the contemporary economics. *Mendeley Data*. Retrieved from: http://dx.doi.org/10.17632/v7dhjnd8xg.2. Access: 16th October 2018.

Maziarz, M. (2019). Methodological pluralism in economics: The 'why' and 'how' of causal inferences. *Filozofia Nauki*, *28*(4), 43–59.

Maziarz, M., & Mróz, R. (2019). Response to Henschen: Causal pluralism in macroeconomics. *Journal of Economic Methodology*, 1–15. DOI: 10.1080/1350178X.2019.1675897

Maziarz, M., & Mróz, R. (2020). A rejoinder to Henschen's response: The issue of VAR and DSGE models. *Journal of Economic Methodology*, First View. DOI: 10.1080/1350178X.2020.1731102

Meek, C., & Glymour, C. (1994). Conditioning and intervening. *The British Journal for the Philosophy of Science*, *45*(4), 1001–1021. DOI: 10.1093/bjps/45.4.1001

Menzies, P., & Price, H. (1993). Causation as a secondary quality. *The British Journal for the Philosophy of Science*, *44*(2), 187–203.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, *13*(2), 151–161. DOI: 10.1080/07350015.1995.10524589

Peace, K., & Chen, D.-G. (2010). *Clinical Trial Methodology*. London: CRC Press.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Oxford: Oxford University Press.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. New York: John Wiley & Sons.

Pearl, J., & Verma, T. S. (1995). A theory of inferred causation. In: *Studies in Logic and the Foundations of Mathematics* (Vol. 134, pp. 789–811). New York: Elsevier.

Pop-Eleches, C. (2006). The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania. *Journal of Political Economy*, *114*(4), 744–773. DOI: 10.1086/506336

Reiss, J. (2005). Causal instrumental variables and interventions. *Philosophy of Science*, *72*(5), 964–976. DOI: 10.1086/508953

Reiss, J. (2013). Philosophy of Economics: A Contemporary Introduction. London: Routledge.

Reiss, J. (2016). *Error in Economics: Towards a More Evidence-Based Methodology*. London: Routledge. DOI: 10.4324/9780203086797

Reiss, J. (2019). Against external validity. *Synthese*, *196*(8), 3103–3121. DOI: 10.1007/s11229-018-1796-6

Reutlinger, A. (2012). Getting rid of interventions. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(4), 787–795. DOI: 10.1016/j.shpsc.2012.05.006

Roth, A. (1988). *Laboratory Experimentation in Economics*. Cambridge: Cambridge University Press.

Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, *12*, 487–508. DOI: 10.1146/annurev.polisci.11.060606.135444

Simon, H. A. (1951). A formal theory of the employment relationship. *Econometrica: Journal of the Econometric Society*, 293–305. DOI: 10.2307/1906815

Simon, H. (1957 [1977]). Causal ordering and identity. In: Simon, H. (ed.) *Models of Man: Social and Rational: Mathematical Essays on Rational Human Behavior in Society Setting* (pp. 55–80). Dordrecht: Springer.

Sims, C. A. (2010). But economics is not an experimental science. *Journal of Economic Perspectives*, *24*(2), 59–68. DOI: 10.1257/jep.24.2.59

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, 72(5), 923–955.

Sprites, P., Glymour, C., & Scheines, R. (1993). *Causation, Prediction and Search*. New York: Springer.

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

Steel, D. (2004). Social mechanisms and causal inference. *Philosophy of the Social Sciences*, *34*(1), 55–78. DOI: 10.1177/0048393103260775

Stegenga, J. (2018). *Medical Nihilism*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198747048.001.0001

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Harvard University Press.

Strevens, M. (2007). Review of Woodward, making things happen. *Philosophy and Phenomenological Research*, 74(1), 233–249.

Strevens, M. (2008). Comments on Woodward, making things happen. *Philosophy and Phenomenological Research*, 77(1), 171–192.

Teira, D., & Reiss, J. (2013). Causality, impartiality and evidence-based policy. In Chao, H. et al. (eds.) *Mechanism and Causality in Biology and Economics* (pp. 207–224). Dordrecht: Springer.

Vogel, H. L., & Werner, R. A. (2015). An analytical review of volatility metrics for bubbles and crashes. *International Review of Financial Analysis*, *38*, 15–28. DOI: 10.1016/j.irfa.2014.11.003

West, K. D. (1988). Bubbles, fads and stock price volatility tests: A partial evaluation. *The Journal of Finance*, *43*(3), 639–656. DOI: 10.1111/j.1540-6261.1988.tb04596.x

Wilde, L. L. (1981). On the use of laboratory experiments in economics. In: *Philosophy in Economics* (pp. 137–148). Dordrecht: Springer.

Woodward, J. F. (2005 [2003]). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press. DOI: 10.1093/0195155270.001.0001

Woodward, J. F. (2012). Agency and interventionist theories. In: Beebee, H. et al. (eds.) *The Oxford Handbook of Causation* (pp. 234–278). Oxford: Oxford University Press.

Woodward, J. F. (2016). Causation and manipulability. In: Zelta, E. (ed.) *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/win2016/entries/causation-mani/

Wright, G. (1971). *Explanation and Understanding*. New York: Cornell University Press.

7 Concluding remarks

Throughout the book, I have studied the methods of causal inference used in contemporary mainstream economics in order to shed light on the meaning of causal conclusions they justify and the limitations of using different types of evidence for policy. My research has been conducted from the perspective of referentialist semantics. According to this philosophical approach, the meaning of words and sentences is given by their referents. Given that economic reality is epistemically inaccessible by means other than by statistical and deductive models, experiments, and still other methods, I take the results of these research methods as referents of the causal claims. In what follows, the meaning of causal claims can be identified by analyzing what types of relations can be uncovered by research methods supporting them.

My research indicates that the use of different research methods in contemporary economics presumes different epistemic views on causality (Section 7.1). Therefore, moderate[1] causal pluralism is the notion of causality *implicitly* accepted by the community of economists. Whereas causal pluralism has received some support in the philosophy of economics and social sciences (e.g., Reiss 2009), my book develops the literature by delivering systematic study of the methods of causal inference and meanings of causal claims. The reconstruction of diverse meanings of causal claims sheds light on how causal evidence can be used for policymaking. Causal claims presupposing different notions of causality have different policy implications. Some policymakers misuse evidence and translate causal claims from the meaning presupposed by economists into some other (e.g., manipulationist) notion. This is unjustified without delivering further empirical support and may lead to unsuccessful interventions (Section 7.2). I conclude the book with the discussion of issues waiting to be resolved (Section7.3).

7.1 Causal pluralism in economics

I have argued that economists, as a group, are moderate causal pluralists. However, the use of specific research methods allows for uncovering different types of relations. These relations, deemed causal, are in agreement with specific definitions of causality present in the philosophical debates. Here, I briefly summarize the mapping of different research methods on the big five philosophical approaches to causality and indicate specific definitions of causality that are the best candidates for being presupposed by the use of particular methods for causal inference.

The regularity view on causality is presupposed by theory-driven econometrics (e.g., the Cowles Commission approach to structural-equation modeling) and cliometric studies of economic history. Importing causal structure from economic theory and conducting econometric research to estimate the values of parameters allows for uncovering empirical regularities that instantiate laws of nature or, employing John Stuart Mill's dictionary, necessary connections. The use of cliometric techniques uncovers spatiotemporal conjunctions of events or empirical regularities, which presupposes the reductionist view on laws (regularity view of laws, or RVL).

The probabilistic view on causality is implicitly accepted by econometricians who inform the direction of causal relation by studying time precedence. The causal interpretation of time series econometrics presupposes a definition of causality that locates itself between Granger's *primafacie* definition and the fully fledged concept based on the knowledge of the whole history of the world. This is the case because econometricians aim at controlling for some confounding factors and include additional variables into their vector autoregressive estimation. However, cross-sectional modeling can also be practiced in an atheoretical way. In such a case, correlational dependencies are interpreted causally on the basis of time precedence. In this case, causes are such variables that modify the conditional distribution of their effects, given some other factors.

The counterfactual view on causality underlies token-level causal claims. Two approaches to making singular claims are at hand. First, economists counterfactually interpret previously established econometric or calibrated theoretical models to establish what would happen if one of the conditions were different. In this way, they establish Galilean counterfactuals. Second, they employ the case study framework to establish singular causal claims and define causes as necessary conditions. The ideal counterfactual causal claim is to deliver evidence for a manipulationist causal claim that addresses the question of what would happen if an intervention were introduced.

Axiomatic, deductive (theoretical) models are used to represent economic mechanisms. The economists that put forward causal claims on the basis of theoretical models implicitly accept the mechanistic view on causality in agreement with New Mechanistic philosophy and Caterina Marchionni's (2017) definition of the economic mechanism. These models are usually models of possible mechanisms, given that there is no empirical support for taking them as the representations of actual mechanisms. However, some calibrated models and DSGE models (the special case of this class) give support for the represented mechanisms to be actual; i.e., operate in the world.

Finally, quasi-experimental research designs and experiments allow for uncovering relations that are invariant under intervention. The former group includes studies that use natural experiments (random assignment to control and treatment groups that are independent of researchers) to infer the effects of a cause.

Experimental studies in economics can be divided into laboratory and field experiments. These designs allow us to infer what are the effects of an intervention introduced by an experimenter and, hence, overlap with Peter Menzies and Hew Price's (1993) agency view: causes can be manipulated to bring about (or raise the probability of) their effects. Given the differentiation of relations that can be uncovered with research methods used by economists to support causal claims, overall, economists are causal pluralists.

In addition to reconstructing the meaning of causal claims, the use of referentialist semantics allows having a glimpse at other presuppositions *implicitly* accepted by economists employing specific methods of causal inference. One of the questions debated in the philosophy of causality is what the relata of causal claims are. Economists seem to be pluralists also in this regard, even though most studies use variables to represent what is under study. Cliometric studies and some econometric research uses variables standing for events. Other econometric models employ variables representing features of phenomena. This gives a hint that causal relations in economics are considered as binding either (types of) events or features of phenomena.

Furthermore, most causal claims in economics are type-level (generic); i.e., they describe the relations between classes of events or similar phenomena rather than single instances. This does not mean that each instance of a cause will produce an effect. On the contrary, virtually all causal claims in economics are of a probabilistic nature and hold only on average. Other (external to model) influences or inherent indeterminism usually make predicting the effect for each instance impossible. The pluralism of epistemic concepts of causality and other philosophical presuppositions implicitly accepted by economists devoted to causal inference have severe implications for using causal claims for policy.

7.2 The meaning of causal claims and translation for policymaking

While the primary purpose of my study is descriptive in nature, analyzing the methods of causal inference and the multitude of epistemic concepts of causality (implicitly) presupposed by these methods allows for throwing light on the uses and misuses of causal evidence in economic policymaking. In particular, I have argued throughout the book that some methods of causal inference are unable to deliver decisive evidence for the causal relation to be invariant under interventions. Policymaking on the basis of insufficient evidence may lead to undertaking careless actions or obtaining unexpected outcomes. Given the framework of referentialist semantics and the focus on reconstructing the meaning of causal claims, I use the notion of 'translation' to denote taking a causal claim and using it in disagreement with what is implied by the evidence behind that claim. For instance, if causal evidence suffices only to put forward a causal claim that is in agreement with the probabilistic view, but a policymaker decides to intervene by changing the relata of causal claims, then they commit themselves to translating the causal claim from the probabilistic understanding into the manipulationist notion. The translation, without additional empirical support, is unjustified.

However, this argument does not imply that only the evidence justifying manipulationist causal claims should be employed as a ground for decisions. I have argued that different types of evidence suffice for different policy actions. It is useful to distinguish the following three types of policymaking. First, some policy actions do not modify the relata of causal claims but use (possibly fallacious) knowledge of causal structure to act in the world. Second, institutional reforms aim at creating a mechanism to promote (but not to warrant) an outcome. For example, policymakers may deliver free access to information in order to improve market efficiency. Finally, 'interventions' rely on modifying causes to influence effects; i.e., they change the relata of causal claims. The three types of policymaking require different classes of causal evidence. These classes allow for inferring causal claims in line with different meanings of 'causality.'

Interventions that modify the values of variables (properties of phenomena) are the type of economic policymaking that is most demanding in regard to causal evidence. Causal claims that justify such interventions need to be invariant under intervention; i.e., presuppose a version of the manipulationist approach. Unfortunately, econometric studies of observational data can only establish causal conclusions on the grounds of an *n*-dimensional econometric model. Policymaking is conducted in a more dimensional world, and therefore excluding the possibility that observed statistical relation results from a common cause (confounding) is impossible. To assert invariance, economists have to uncover the actual causal structure and obtain knowledge that an intervention under interest does not change it. I have argued that, given the limitations of observational studies, this can be done only by means of experimentation or (design-based) studies that mimic the gold standard of causal inference and allow for putting forth causal claims presupposing the manipulationist view on causality. Therefore, if a policymaker takes a causal claim presupposing other views on causality and translates it into the manipulationist notion in order to justify their intervention, they misuse causal evidence.

Institutional reforms require having knowledge of how a mechanism operates. Mechanistic evidence is unable to warrant that causal claims are invariant (i.e., that actions on causes produce expected effects) because most models represent single mechanisms, while, in the economy, several mechanisms operate at the same time. Furthermore, economists put forward causal claims describing mechanisms on the basis of models of possible mechanisms. Mechanistic evidence can be used as justification for institutional reforms relying on modifying the factors that shape preferences and decisions of economic agents with a view to promote (but not warrant) the emergence of mechanisms represented by models.

Policy actions can be based on partial knowledge of causal structure. In fact, even knowledge of spurious causal relations (e.g., produced by a common cause) is sufficient because policy actions do not interfere with how phenomena of interest are produced by causal factors. Therefore, the studies presupposing probabilistic or regularity views on causality are sufficient for this type of policymaking. The prime example is using the color of one's car to predict the likelihood of accidents. Even though it is not the redness of the car but some characteristics of car owner (common cause) that cause both reckless driving and preference for red cars. Insurance companies commit to the common-cause fallacy on a daily basis but are still able to successfully act in the world. Another example is to use the knowledge of (potentially spurious) relation between public debt and economic growth to locate a factory in a country of high growth potential. Such (spurious) correlations can be used for policy as long as policymakers' actions do not interfere with the actual causal structure that create the phenomena of interest.

7.3 Further research

Philosophy and economics have in common being interested in causality since their beginnings. In the case of former, these beginnings can be traced back to pre-Socratic thought (Andriopoulos 1995). In the case of the latter, causal language appears even in the title of Adam Smith's (1950 [1776]) book. However, more questions are waiting to be addressed at the intersection of philosophy of causality and causal inference in economics.

First, when conducting the systematic literature review, I have found many cases of pairs of similar studies differing only in regard to the language used to formulate conclusions: while one study is explicitly causal, the other uses correlational and associational language. I hope that studying such cases could indicate the differences between contexts of analyses that allow for formulating causal conclusions only in one case.

Second, causal pluralism is adequate to all economists as the epistemic view. But this raises the question of what are the views of each economist? Considering that contemporary economics is so advanced that each researcher specializes in one or a few related research methods, one can hypothesize that each economist accepts as an epistemic definition of causality the view that can be uncovered in this way. On a related note, the views on causality and causal evidence held by policymakers (also those responsible for economic policy at national and international levels) stay unknown. While my book fills this gap in the literature, further research – and, especially, descriptive studies – are needed to inform the relatively new field of philosophy of economics; i.e., the studies on the use of causal knowledge (cf. Cartwright and Hardie 2012). Third, my purpose has been to uncover what are the epistemic views of economists on causality. In other words, I have aimed at uncovering what types of relations are considered as causal by economists. However, economists (both as a group and as single researchers) can accept different epistemic and ontological stances. For instance, it is possible to accept correlational results as an indication of causality (epistemic definition) but acknowledge that causal relations are invariant under intervention or are produced by mechanisms (ontological view).

Finally, I have encountered many cases of inconsistent results in the sample under review. Approximately 5% of econometric studies putting forward explicitly causal conclusions published by three top economic journals (*American Economic Review*, *Journal of Political Economy*, and *Quarterly Journal of Economics*) contradict other studies in the sample. The phenomenon of inconsistent results (cf. Goldfarb 1997) has detrimental effects on making policy-relevant inferences from literature reviews and informing theoretical discourse. However, inconsistent results are also widespread in contemporary economics: given that my sample covers only a small fraction of journals publishing econometric research, the 5% estimate is the lower bound and the number of studies contradicting other results is likely to be much higher. Today, we have insufficient knowledge of the causes of why inconsistent econometric studies appear, but the urgent question is how to make inferences from strains of literature including inconsistent results. I hope that tackling these and other questions will not only develop our understanding of causality and, especially, causal inferences in the sciences, but also improve the quality of policymaking with benefits for society.

Note

1 I support causal pluralism understood as a stance accepting several different definitions of causality. This view is 'moderate' in comparison to radical pluralism accepting that each use of causal language implicitly accepts a (slightly) different notion of causality (cf. Maziarz and Mróz 2019).

Bibliography

Andriopoulos, D. Z. (1995). Concepts of causality in pre-socratic philosophy. In: *Natural Sciences and Human Thought* (pp. 93–99). Berlin: Springer.

Cartwright, N., & Hardie, J. (2012). *Evidence-Based Policy: A Practical Guide to Doing It Better.* Oxford: Oxford University Press. DOI: 10.1093/acprof:osobl/9780199841608.001.0001

Goldfarb, R. S. (1997). Now you see it, now you don't: Emerging contrary results in economics. *Journal of Economic Methodology*, *4*(2), 221–244. DOI: 10.1080/13501789700000016

Marchionni, C. (2017). Mechanisms in economics. In: Glennan, S. & Illari, Ph. (eds.) *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. London: Routledge.

Maziarz, M., & Mróz, R. (2019). Response to Henschen: Causal pluralism in macroeconomics. *Journal of Economic Methodology*, 1–15. DOI: 10.1080/1350178X.2019.1675897

Menzies, P., & Price, H. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science*, *44*, 187–187.

Reiss, J. (2009). Causation in the social sciences: Evidence, inference, and purpose. *Philosophy of the Social Sciences*, *39*(1), 20–40. DOI: 10.1177/0048393108328150

Smith, A. (1950 [1776]). *An Inquiry Into the Nature and Causes of the Wealth of Nations.* Methuen. Retrieved from: website.aub.lb/fas/Documents/Adam-SMITH.doc

Vandenbroucke, J., Broadbent, A., & Pearce, N. (2016). Causality and causal inference in epidemiology: The need for a pluralistic approach. *International Journal of Epidemiology*, 1776–1786. DOI: 10.1093/ije/dyv341

Index

Note: Page numbers in italic indicate a figure on the corresponding page.