



Aleksander Ostapiuk

The Eclipse of Value-Free Economics

The concept of multiple self
versus homo economicus



Publishing House of Wroclaw University of Economics and Business

The Eclipse of Value-Free Economics

**The concept of multiple self
versus homo economicus**

Aleksander Ostapiuk

The Eclipse of Value-Free Economics

**The concept of multiple self
versus homo economicus**



Publishing House of Wrocław University of Economics and Business
Wrocław 2020

Reviewer

Łukasz Hardt

Copy-editing

Agnieszka Flasińska

Layout

Barbara Łopusiewicz

Proof reading

Barbara Łopusiewicz

Typesetting

Małgorzata Myszkowska

Cover design

Beata Dębska

On the cover, a fragment of a painting by Hieronymus Bosch,
The Garden of Earthly Delights (Public Domain)

All rights reserved. No part of this book may be reproduced in any form
or in any means without the prior written permission of the Publisher

© Copyright by Publishing House of Wrocław University of Economics and Business
Wrocław 2020

ISBN 978-83-7695-853-8

Printing: TOTEM

Contents

| | |
|--|------------|
| Introduction | 7 |
| 1. The origins of value-free economics | 17 |
| 1.1. Introduction | 17 |
| 1.2. Pareto turn – economics and psychology | 19 |
| 1.3. Robbins and value-free economics | 22 |
| 1.4. Positive versus normative approach | 27 |
| 1.5. Conclusions..... | 32 |
| 2. Lakatosian perspective on Becker’s economic approach | 34 |
| 2.1. Introduction | 34 |
| 2.2. Becker’s economic approach as scientific research programme | 38 |
| 2.3. Becker’s methodology | 41 |
| 2.4. Problems with instrumental rationality and utility maximization | 44 |
| 2.5. Becker’s unsuccessful escape from tautology | 52 |
| 2.6. Conclusions..... | 55 |
| 3. Economics of happiness..... | 59 |
| 3.1. Introduction | 59 |
| 3.2. Subjective well-being | 63 |
| 3.3. The criticism of revealed preference theory | 67 |
| 3.4. Economics of happiness – general insights..... | 77 |
| 3.5. Objective well-being approaches | 78 |
| 3.6. Conclusions | 87 |
| 4. The conception of multiple self. Reason versus emotions | 90 |
| 4.1. Introduction | 90 |
| 4.2. Personal identity and economics | 90 |
| 4.3. The conception of multiple self..... | 100 |
| 4.4. The conception of multiple self over time..... | 103 |
| 4.5. Weakness of will | 106 |
| 4.6. Conclusions..... | 111 |
| 5. Philosophy of science and value-free economics | 113 |
| 5.1. Introduction | 113 |
| 5.2. Positivist methodology of economics | 121 |

| | | |
|-----------|---|------------|
| 5.3. | Weber and value-free economics..... | 126 |
| 5.4. | Consequentialism versus deontology..... | 130 |
| 6. | The capability approach..... | 137 |
| 6.1. | Introduction..... | 137 |
| 6.2. | The criticism of neoclassical economics..... | 137 |
| 6.3. | Functionings and capabilities..... | 140 |
| 6.4. | Freedom as an intrinsic value..... | 143 |
| 6.5. | Reason and autonomy..... | 147 |
| 6.6. | The criticism of the capability approach..... | 153 |
| 6.7. | Conclusions..... | 157 |
| 7. | Libertarian paternalism. The discussion on freedom and well-being..... | 159 |
| 7.1. | Introduction..... | 159 |
| 7.2. | Psychological background..... | 161 |
| 7.3. | Libertarian paternalism..... | 163 |
| 7.3.1. | Introduction..... | 163 |
| 7.3.2. | Choice architecture..... | 164 |
| 7.3.3. | Default option versus freedom..... | 168 |
| 7.3.4. | Some constraints on libertarian paternalism..... | 170 |
| 7.4. | Libertarian paternalism in the light of capability approach..... | 171 |
| 7.5. | The criticism of libertarian paternalism..... | 172 |
| 7.5.1. | Introduction..... | 172 |
| 7.5.2. | The concept of purified preferences. The problems with the inner agent..... | 173 |
| 7.5.3. | Paternalism is not inevitable..... | 183 |
| 7.5.4. | Libertarian paternalism does not provide freedom..... | 186 |
| 7.6. | Conclusions..... | 193 |
| | Conclusions..... | 195 |
| | References..... | 202 |

Introduction

In the last years the criticism of neoclassical economics together with the conception of *homo economicus* has been increasing. Of course, neoclassical economists have been criticized from the very beginning. However, recently the criticism has intensified. It comes from within economics, and from outside economics. Neoclassical economics is mostly criticized by other social sciences (e.g., psychology, sociology, political sciences) and philosophy. The most visible strand of criticism is descriptive and it concerns *homo economicus* (e.g., people are not: 1) fully rational, 2) fully egoistic, 3) atoms independent of society). The criticism also concerns revealed preference theory and the assumption that people know what is best for them. Beyond descriptive criticism, neoclassical economics also faces more sophisticated criticism from philosophy. The big part of criticism concerns a view cherished by many economists that economics is value-free. This belief is strong among some economists who think that value-freeness is what distinguishes economics from the other social sciences and makes economics closer to hard sciences. A methodological criticism concerns also the method used by economists, e.g., deduction, the overuse of mathematics, overreliance on models, idealizations and focus on predictive power.

When we analyse how strongly neoclassical economics has been criticized in the last years, it comes as no surprise that not many admit being diehard neoclassical economists. In some circles it is a dirty word and not many people want to be connected with it. The result of this extensive criticism is the postulate that economics should change. Namely, it should be more pluralistic and open for heterodoxy. Of course, different schools propose their own visions of economics. Behavioural economists want economics to incorporate psychology, institutional economists want economics to incorporate sociology, and philosophers of economics want economics to incorporate philosophy.

Some economists argue that economics has changed due to the incorporation of the knowledge from other social sciences like psychology, sociology, etc. (e.g. (Cedrini, & Fontana, 2018; Colander, 2000, 2010; Colander, Holt, & Rosser, 2004, 2011; Davis, 2006, 2008; Fontana, 2014; Vromen, 2007)). The most telling case is behavioural economics which started to be perceived as a new mainstream (Angner, 2019; Thaler, 2016; Truc, 2018). Moreover, it is argued that economics is not so heavily theoretical anymore and we could observe 'empirical turn' in the last years (Backhouse, & Cherrier, 2017). In general, the process of opening

economics to other social sciences has been noticed. It is called “reverse imperialism.”

However, many economists (the author of this book among them) argue that despite the changes, the foundations of economics have not really changed and mainstream economics predominately works as it used to be (e.g., (Berg, & Gigerenzer, 2010; Dobusch, & Kapeller, 2009; Dow, 2012; Elsner, 2013; Fine, 2019; Fourcade, Ollion, & Algan, 2015; Hodgson, 2012b; Madra, 2016; Ostapiuk, 2019a)). Everyone agrees that economics has incorporated many insights from other social sciences. However, they are only adjustments and the paradigm has not changed. Economists still rely on rationality, utility, equilibrium, *homo economicus*, revealed preferences, deductive thinking, and mathematics.

To recognize that neoclassical economics has not changed we can look at the syllabuses and textbooks used by most economics students around the world (Colander, 2015; Graupe, 2019). Of course, the change in textbooks and syllabuses is slower than the changes in specialistic journals. However, they have an enormous influence on millions of economics students. As Samuelson writes “I don’t care who writes a nation’s laws – or crafts its advanced treaties – if I can write its economics textbooks” (Samuelson, as cited in Skousen, 1997, p. 150). The unmodified economics textbooks show that the change concerns particular branches of economics (e.g., behavioural economics, institutional economics, evolutionary economics, complexity economics), whereas most economists use a toolkit of neoclassical economics. What can be even more important in this story is how economists are perceived by other social scientists and public opinion. Here, we will not go into the discussion of how much blame we should assess to economists for the bad image of *homo economicus* which is perceived as a fully egoistic and rational calculator. Now, it is important to point out that the criticism of neoclassical economics is not a straw man. Some economists really believe that *homo economicus* exists and that economics is value-free. It is especially important, because theories are the lenses by which economists look at reality and people.

The introduction of this book should give a reader an impression, shared by many, that economics is in a process of some kind of “paradigm shift” ((Kuhn, 1962); see also (Colander et al., 2004; Davis, 2006; Dobusch, & Kapeller, 2012; Dow, 2012; Etzioni, 2011; Thaler, 2015, 2016)).¹

Most economists see the limitations of neoclassical economics. However, they are not sure how the future of economics should look like. Whether neoclassical economics should change dramatically or adjust a little by adding the bits of

¹ The expectation for change in economics is mostly visible in the case of movements like “Institute for New Economic Thinking” and “Rethinking economics”.

psychology, sociology, etc. This is the conundrum that this book addresses. The main question is: Do the weaknesses of value-free economics imply the need for a paradigm shift?

As was indicated earlier, despite the criticism of value-free economics, a paradigm (in Kuhnian understanding (Kuhn, 1962)) in economics has not changed. There seem to be two main reasons for this. Firstly, the criticism is not severe and structured enough. Secondly, the alternative paradigms proposed by the critics are not conceptually sound enough. To answer the first objection the author synthesizes criticisms from different perspectives (from behavioural to methodological). Special attention is put on choices over time, because it seems that in this area value-free economics has the most problems (methodological and descriptive). The enriched concept of multiple self is proposed and investigated. Thanks to that, it is possible to identify the main weaknesses of value-free economics in order to overcome them. However, it is not enough to present the criticisms towards value-free economics. For scientists, a bad paradigm is better than no paradigm (Kuhn, 1962). This is why the author considers whether value-based and goal-oriented economics with normative approaches such as economics of happiness, capability approach, libertarian paternalism, and the concept of multiple self can be the alternative paradigm for value-free economics.

What a reader can gather from the introduction so far, the scope of this book is quite extensive. It is impossible to fully analyse economics as science on 200 pages. Therefore, the focus is put on a few main ideas which can contribute to the current debate on economics. Firstly, I use ‘scientific research programme’ proposed by Lakatos (1980). It is quite an old-fashioned approach that has fallen from grace in economics. However, it is a very useful method to analyse how neoclassical economics works. I indicate and analyse ‘hard core’ and ‘protective belt.’ Thanks to this analysis we can understand how economists can absorb the descriptive criticism and leave hard core and paradigm of neoclassical economics unchanged. Tautologicality of rationality and utility used in economics is well-known (Boland, 1981; Hodgson, 2012b; Kirchgässner, 2013; Sen, 1977). However, Lakatosian method explains how neoclassical economists could absorb descriptive criticism towards *homo economicus* by changing the very meaning of the terms ‘rationality’ and ‘utility.’ Lakatosian approach sets the stage for the whole book. It concerns the methodology of economics, because the descriptive criticism is not enough to change neoclassical economics which is based on axioms. Therefore, we need to analyse the methodological assumption of neoclassical economics from which the most important one is its value-freeness.

At this point, I used a notion of value-freeness a few times. Moreover, the title of the book contains “value-free economics.” Now I will give only a few words of explanation but what I mean by value-free economics should be fully understandable after reading this book. In general, I treat value-freeness as one

of the most important assumptions of neoclassical economics, which in many ways dictates how economists should do their work. Firstly, it is assumed that economists are like engineers who only try to solve a problem but they do not analyse the problem itself. It means that economics is all about means not ends. To use the most popular definition, “economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses” (Robbins, 1932, p. 16). It means that economists do not decide what is good or bad and what goals people should have. The only thing that matters is effectiveness. Economists want to be objective scientists like physicists (physics envy). Therefore, economists do not want to deal with normative statements, values, and goals. That is why economists rely on axiomatic revealed preference theory which is perceived as a neutral framework. This framework is similar to mathematics in which we have axioms from which we deduce theories. The belief in objectivity is supported by the axiomatic assumption of revealed preference theory – what people do is the best for them. In economists’ eyes it is not a normative statement – we just wait and see what people choose; and what people choose is the best for them. Otherwise why they would choose something? Therefore, many economists think that they do not need to be entangled in a discussion on human goals and motivations. Neoclassical economists rely on human choices (revealed preferences) (Gul, & Pesendorfer, 2008).

The next part of the title concerns the conception of multiple self. It must be stated once more that the entire criticism of neoclassical economics with *homo economicus* is so vast that a separate book should be devoted to this issue (see (Dzionek-Kozłowska, 2018)). Therefore, here the focus is put on choices over time. There we can observe the crucial problems with neoclassical economics. It especially concerns revealed preference theory with its assumptions: people choose what is the best for them, they have stable preferences and can compare choices over time by utility measurements.

The axiomatic assumption that what people do is the best for them leads to many problems. When economists do not decide what is the best for people and they assume that people choose what is the best for them, they support ‘short-term human’ who frequently falls into temptation and chooses simple pleasures and regrets it later. Revealed preference theory provides economics with false objectivity (assuming that what people do is the best for them is a normative claim) and it has a negative impact on individuals’ well-being.

The conception of multiple self shows that people have different preferences over time which cannot be integrated within utilitarian measurement. The conception of multiple self is used as a starting point for the discussion on the problems of revealed preference theory. The author develops this idea and he proposes the enriched conception of multiple self with the conceptual framework designed to normatively decide between selves (short-term versus long-term). In the

author's opinion value-free economics do not have a luxury not to decide what is good for people (not deciding means supporting "short-term human").

Even though the focus of the book is on choices over time, the author also synthesizes different criticisms toward neoclassical economics which aspires to be value-free. Both kinds of criticisms are investigated: descriptive and methodological. After the synthesis of criticism toward value-free economics, the author investigates some alternative normative approaches in economics. The goal is to check whether normative approaches can be a new paradigm that changes how economics is practiced. It should answer the question how economics should look in the future. Trying to answer such a broad question is essential, because it is not enough to criticize an old paradigm. To have a successful 'paradigm shift' scientists need to have a better paradigm than the old one. The author analyses some normative approaches, e.g., economics of happiness, the conception of multiple self, capability approach, and libertarian paternalism. The common theme in these approaches is that they ask what is good for people and how to improve people's life. In some way it is the same kind of questions that economists ask at the beginning (e.g., (Marshall, 1920; Mill, 1874; Smith, 1869)). The problem with these normative approaches is if it is possible to improve people's life and guarantee their freedom at the same time.

I think it is a good idea to tell the readers what they can expect from the book. It is also crucial to indicate to whom the book is written and what the author wants to achieve. This book is interdisciplinary in its approach. Many topics are analysed (philosophy, economics, psychology, etc.) and the author connects ideas from different fields. However, philosophy of economics is the main perspective which is used in this book (1) rationality, 2) ethics, 3) methodology). The book is written predominately for people interested in the philosophy and methodology of economics. However, the broad scope of the book means that it is not solely written for economists but for a wider audience. It can be an engaging read (hopefully!) for people interested in economics as a science and also for scientists who are curious about the responsibilities and goals of social sciences. Of course, economics differs from other social sciences. However, the questions on positive vs normative dichotomy are as important in psychology, sociology, and political science.

The main goal of this book is to present the argument that economics cannot be value-free, must be open toward methodological pluralism and normative approaches. On their own, these statements are not new. They had been made by many economists who are more experienced than the author of the book. However, Lakatosian framework together with multiple self conception used here can form a new argument that can persuade economists who still are not sure whether economists should deal with normative approaches, because they are afraid to be unscientific.

The book engages in a general discussion on ‘paradigm shift’ which is an extremely rich and large topic. Even though the scope of the book is quite extensive, the main goal is always underlined. It is to present the limits of neoclassical economics that wants to be value-free. In the book I argue that being completely value-free is impossible and this belief leads economists to methodological and practical problems. The argument presented in this book is as follows. If science cannot be completely value-free, economists should put their values on the table to analyse them and make them transparent.

Some claims made in this book can be controversial for the orthodox economists, because the very method they used is criticized. However, for me a possible discussion is good news. Starting the ‘conversation’ (Rorty, 1979) is one of the goals of this book. This is a conversation, because there are no ironclad arguments about how economics should be done. Therefore, I do not want to pretend that I have arguments that will end “strife over methods” (*Methodenstreit*) which seems to be a never-ending story. However, at least we can discuss our methodological positions. The message of this book is that value-free economics uses normative values that are hidden in the disguise of objectivity. Unfortunately, many economists do not see this. This book should show why this is a crucial problem and why economists should not believe that they are value-free engineers.

Now, let us have a brief presentation of the content of this book. In **Chapter 1** the origins of neoclassical economics and the process of becoming value-free science are presented. The main assumptions of neoclassical economics are investigated. Special attention is paid to revealed preference theory and positive/normative dichotomy. Thanks to the historical analysis a reader should understand what the author means by value-free economics.

In **Chapter 2** ‘scientific research programme’ (Lakatos) is used to analyse Becker’s economic approach. The economic approach is chosen, because it is treated as the embodiment of neoclassical economics. The conclusions from this analysis are extrapolated on value-free economics. The ‘hard core’ is indicated (instrumental rationality and utility *ad libitum*) together with ‘protective belt’ and ‘negative heuristics.’ It is presented how thanks to ‘protective belt’ economists could absorb descriptive criticism and ‘hard core’ has remained unchanged. It is argued that the economic approach and value-free economics are degenerative programmes, because instrumental rationality and utility *ad libitum* lead to tautology. The analysis of value-free economics by ‘scientific research programme’ shows that descriptive criticism of *homo economicus* (e.g. people are altruistic and irrational) is not enough to change the ‘hard core’ of value-free economics. However, due to ‘scientific research programme’ we obtain the tools to see the problems and limitations of the paradigm used by neoclassical economists.

If we agree that value-free economics is a degenerative programme, then the natural question which arises is what economists should use instead. The answer to this question is the topic of **Chapter 3**. The natural alternative is economics of happiness which tries to answer the question what is good for people and what improves their well-being. The main assumptions and origins of economics of happiness are presented. Moreover, the limits of this approach are demonstrated (1) malleability of preferences, 2) other values than happiness, e.g., freedom). Due to these problems the other approaches that indicate objective goals are analysed. The comprehensive analysis of economics of happiness can indicate the strengths and weaknesses of this approach. Thanks to that, it is possible to answer the question whether economics of happiness is better (both descriptively and methodologically) than revealed preference theory when it comes to well-being.

Chapter 4 starts with a discussion on personal identity. Firstly, the economics' perception of self over time is analysed. Secondly, the philosophical discussion which concerns the question whether it is possible to identify one person over time is investigated (Hume, Lock, Parfit). Then, we move into the analysis of the conception of multiple self which is critical towards neoclassical economics, because the latter copes with different choices over time within the utilitarian framework (the discounted utility model (Samuelson, 1937)). The conception of multiple self demonstrates that people do not always choose what is best for them which is a core of revealed preference theory. To support this claim the philosophical discussion on weakness of will is analysed within the context of value-free economics. The main goal of this chapter is to provide a normative framework thanks to which we can decide between selves ('short-term human' vs 'long-term human'). It has not been done successfully by the inventors of multiple self conception (Elster, Schelling). Moreover, the enriched conception of multiple self can indicate what people's real preferences are.

The conception of multiple self is normative, because we try to establish goals and values that are independent of people's choices. As was indicated before, many economists do not like normativity and treat this endeavour as unscientific. In **Chapter 5** the author delves into the epistemological discussion on the value-freeness of science to show that this perception of normative approaches is not justified. Over the last 50 years, we can observe the criticism towards the idea of objective science and the possibility of finding the Truth² and iron-clad methodology (see (Feyerabend, 1993; Kuhn, 1962; Lakatos, 1980; McCloskey, 1985; Quine, 1969; Rorty, 1979)). The conclusions from the philosophy of science have not been accepted by many economists who are under the spell of logical positivism or some versions of it (Blaug, 1992; Caldwell, 1982; Fiedor,

² Truth with capital T refers to the epistemological ideal of absolute and objective Truth.

& Ostapiuk, 2017; Hands, 2001; Hardt, 2017; Hausman, 2007; Hausman, McPherson, & Satz, 2017; Lawson, 1997, 2003; McCloskey, 1983, 1985; Reiss, 2013). The main unresolved problem is a normative/positive distinction which is still a cherished methodological position in orthodox economics (Dasgupta, 2005; Fiedor, & Ostapiuk, 2017; Friedman, 1953; Hausman et al. 2017; Lawson, 2003, 2015; Mirowski, 1991, 2002; Ostapiuk, 2017a, 2019a; Reiss, 2017; Zboroń, 2013). The methodological literature identifying the problems with positive/normative distinction is still not well-known to many economists (Blaug, 1992; Colander, & Su, 2015; Czarny, 2010; Fiedor, 2013; Hands, 2012; Kincaid, Dupré, & Wylie, 2007; Kwarciański, 2016; Mongin, 2006; Myrdal, 1970; Putnam, 2002; Putnam, & Walsh, 2011; Sedláček, 2011) and it concerns mainly philosophers of science and philosophers of economics. Therefore, **Chapter 5** synthesizes and clarifies this large discussion in the context of value-free economics.

Next in **Chapter 5**, Max Weber is analysed to show that many economists misinterpreted his concepts of value-freeness and ideal type. He is commonly used as a justification for value-free economics. It is argued that it was not what he meant. In the end, Weber argued that scientists cannot escape from using values in their research (Weber, 2012). At the end of **Chapter 5** two ethical systems (consequentialism and deontology) are analysed. The goal is not to compare these two ethical systems but to present and argue for a difference between intrinsic and instrumental values. Intrinsic values are independent of consequences, and their existence is necessary if we want to choose which self should be in control. Kant's deontological system is presented as an attempt to establish the ultimate intrinsic values. It shows the limits of consequential reasoning used in economics (revealed preference theory). Moreover, the analysis of deontology is used to establish goals that are independent from people's choices.

The crucial problem with establishing intrinsic goals that are independent of people's choices is that they can curb people's sovereignty and freedom, because revealed preferences are dismissed. In **Chapter 6** the capability approach is analysed to reconcile objective goals with freedom. Sen's capability approach is also analysed, because he is the most salient critic of neoclassical economics and revealed preference theory. Moreover, capability approach is used to support the argument made in this book that reason establishes freedom and autonomy. Sen's approach is connected and analysed within the context of Kant and contemporary philosophers (Dworkin and Frankfurt). The argument is that acting on given preferences as in neoclassical economics is not sufficient to establish freedom. Besides analysing capability approach, the criticism towards Sen's approach is presented. Especially the argument that his approach is paternalistic. The difference between Sen's and Nussbaum's account is presented to show why Sen's approach perceives freedom as intrinsic value. His framework is chosen to support the conception of 'long-term human.'

First six chapters concern methodology. They conclude that economics cannot be value-free. Therefore, economists need to ‘get their hands dirty’ and be entangled in the normative discussion on human well-being. **Chapter 7** concerns libertarian paternalism and is presented as a practical application of the previous methodological chapters. Moreover, the conception of libertarian paternalism proposed by Thaler and Sunstein (2008) has attracted a lot of attention, mostly negative one (Binder, 2014; Fumagalli, 2016; Gigerenzer, 2015; Glaeser, 2006; Grüne-Yanoff, 2012; Hédoïn, 2017; Infante, Lecouteux, & Sugden, 2016; McQuillin, & Sugden, 2012; Mitchell, 2005; Qizilbash, 2009, 2012; Schnellenbach, 2016; Sugden, 2008b, 2009a, 2010, 2015, 2017, 2018; White, 2013). One of the reasons for such strong criticism is Thaler’s and Sunstein’s inability to base libertarian paternalism upon a strong philosophical foundation. They claim that libertarian paternalism guarantees both freedom and well-being for individuals. However, they were not able to provide enough arguments to make this claim convincing for many.

I use the conclusions from this book to put libertarian paternalism on a more solid methodological foundation. The discussion is not solely focused on libertarian paternalism but also on behavioural welfare economists that cast doubts on consumer sovereignty. They argue that we can find purified preferences which rational human would choose to have if it were not for cognitive biases. This claim is strongly criticized. To answer some of the criticisms I use the concept of ‘long-term human’ to argue that we can find, at least in some cases, purified preferences. Even though libertarian paternalism can be defended against a lot of criticism, it does not mean that we should close our eyes to the dangers of libertarian paternalism. Therefore, only a limited version of libertarian paternalism is supported.

In the end, it is presented that we cannot discuss concepts like libertarian paternalism without indicating our values first. This idea resonates with the main goal of this book – economists need to put their values on the table and make them transparent. They should not be hidden in the disguise of objectivity. Of course, it can mean that we will never end this debate but at least our ‘conversation’ can be more fruitful.

Chapter 1

The origins of value-free economics

Economic theorists, like French chefs in regard to food, have developed stylized models whose ingredients are limited by some unwritten rules. Just as traditional French cooking does not use seaweed or raw fish, so neoclassical models do not make assumptions derived from psychology, anthropology, or sociology

(Akerlof, 1984, p. 2)

1.1. Introduction

We are on the verge of a 'paradigm shift'. *Homo economicus* and the assumptions of neoclassical economics have been criticized for a long time. However, it seems that fundamental change still has not come and neoclassical economics prevails as a dominant paradigm, despite reverse imperialism. One of the reasons may be that many economists still believe in positive-normative dichotomy and they think that economics must be value-free. The economic approach is the example where this faith can lead. Becker's economic approach is analysed as the embodiment and capstone of value-free economics (Becker, 1968, 1974, 1976, 1993; Becker, & Murphy, 1988; Stigler, & Becker, 1977). It can be perceived as the lens by which economists look at reality. Many economists believe in 'hard core' assumptions of the economic approach (maximization of utility and instrumental rationality). Becker's approach is so attractive for economics, because it is an extremely effective research programme that can absorb every kind of criticisms and it has made economics the most influential social science (Lazear, 2000; Mäki, 2009a; Mäki, Walsh, & Pinto, 2017). The economic approach is analysed in the context of 'scientific research programme' (Lakatos, 1980). The author examines 'hard core' of economic approach (maximization of utility, instrumental rationality), 'protective belt' and 'heuristics.' It is checked whether economic approach is 'degenerative programme.' The conclusions from this analysis can be extrapolated on value-free economics, because in the author's opinion the economic approach in some respects can be perceived as the embodiment of value-free economics.

The analysis of Becker's economic approach is so crucial, because it gives one the possibility to understand what is the 'hard core' of neoclassical/value-

-free economics.¹ Still a lot of neoclassical economists perceive utility and instrumental rationality as positive conceptions. Positive/normative dichotomy is investigated to assess whether economists' belief is justified and to what problems it leads. After reading the next two chapters, it will be possible to decide whether economics should stick to an effective value-free approach or should it look towards some goal-oriented normative approaches.

Before moving to Becker and his economic approach it is necessary to analyse the history of economic thought, which is the goal of the first chapter. It will enable us to indicate what are the cornerstones and the origins of neoclassical economics, and how they have changed over the years. We need to begin with the marginal revolution which started in the early 1870s when three independent economists (Jevons, Menger, and Walras) came up with a similar idea. The notion of revolution is not accidental, because the primary interest of economics shifted at that time from the production of wealth to utility which is gained from consumption. Moreover, the focus shifted from analysing the whole economy towards individuals and their wants (methodological individualism).

It is hard to overestimate the influence of utility conception on value-free economics. In this context, it can even be stated that Bentham was one of the fathers of contemporary economics. Certainly he is regarded as the founder of utilitarianism. For him utility can mean everything that is beneficial for someone. What is more important, in Bentham's opinion people always seek to maximize utility and this is because "nature has placed mankind under the governance of two sovereign masters, pain and pleasure" (Bentham, 1907, I.1). Furthermore, Bentham underlines that pleasure and pain are not only our empirical 'masters' but also they define how we should behave "It is for them alone to point out what we ought to do, as well as to determine what we shall do" (Bentham, 1907, I.1). This accentuation is essential, because in economics utilitarianism is discerned only as a system that describes human behaviour. It deals with individual choices not with ethics in which a predominant feature is universalism.²

¹ The term 'neoclassical economics' is often overused and it leads to obscurity (Colander, 2000). In this book terms 'neoclassical' and 'value-free economics' are used interchangeably, because both of them are based mainly on revealed preference theory which is not interested in human goals, assumes that people do what is best for them, and rationality is defined by technical criteria (consistency, completeness of preferences)

² Universalism means that ethical systems have some rules that must be followed in all circumstances by all people. Generally, universalism is identified with deontological approaches (Kant's categorical imperative) when people must follow some duty/rules no matter what the consequences are (Kant, 1993). However, also consequentialist approaches (utilitarianism) are universal in the sense that all people need to follow some rule (maximize utility). More about the relationship between ethics and economics see (Etzioni; 1988; Fiedor, & Ostapiuk, 2017; Hardt, 2020; Róna, & Zsolnai, 2017).

Especially Jevons relied on utilitarianism. For him all pleasures and pains can be reduced to a single dimension of utility and all desires can be simplified to the desire to maximize utility. In the first generation of neoclassical theory, the utility-based choice can be characterized by the cardinal and hedonistic notion of utility. Cardinal approach means that it is possible to assess utility and give it numerical values. The hedonistic notion is connected with the amount of pleasurable or painful psychical feeling that consumer gets from some goods or actions (Hands, 2009). What makes value-free economics and Bentham so similar is their reluctance to indicate some goals of human actions. Bentham finds the maximization of utility as an end itself. Therefore, he criticizes the Greek ancient philosophers who believed that some pleasures depreciate human being. The English philosopher scoffs at their searching for the highest good (*summum bonum*) and claims that this is the “consummate nonsense” (Bentham, 1983, p. 134). In his opinion, searching for something better and different from pleasure does not have any sense, as it is impossible to be found. Plato, with his realm of ideas, is for Bentham “the master manufacturer of nonsense” (Bentham, 1983, p. 137). Bentham also states that there is no distinction between better and worse pleasures. He writes: “Prejudice apart, the game of push-pin is of equal value with the arts and sciences of music and poetry” (Bentham, 1825, p. 206). In his opinion, both activities are equally good if they give the same amount of utility. However, there is a crucial difference between not deciding which goods are more worthy (Bentham) and assuming that people always maximize their utility which implies that people choose what is the best for them (value-free economics).

1.2. Pareto turn – economics and psychology

At the beginning of the 20th century, economics got rid of psychology and this process is called ‘Pareto turn.’ Many economists, and among them Bruni and Sugden (2007), comprehensively explained how it was done. Therefore, in this section only the main factors in this process will be analysed. To begin with, it must be stated that marginalists believed in cardinal utility. However, they could not find a convincing method to measure utility (see (Stigler, 1950)). Of course, the lack of a method to measure utility attracted a lot of criticism. Robbins underlines that marginal utility theory was severely criticized, because of its psychological hedonism foundation:

The borderlands of Economics are the happy hunting ground of minds averse to the effort of exact thought, and, in these ambiguous regions, in recent years, endless time has been devoted to attacks on the alleged psychological assumptions of Economic Science. (Robbins, 1932, p. 83)

Psychological hedonism quickly lost intellectual credibility and it was evident for marginalists that they had to reform its theoretical foundations. Cardinal

approach with the belief that it is possible to assess utility was untenable. Therefore, economists started to use the ordinalistic approach where consumers can only rank their preferences. It means that consumers may not be able to indicate the exact amounts of utilities that they derive from commodities, but they can judge whether the satisfaction obtained from a commodity is equal, lower than, or higher than another. In this perspective, consumers are presented with a number of various combinations of goods and they can rank or order them according to the personal scale of their preferences. The inability to assess utility not only resulted in a new ordinal approach but also in escaping from a debate on human goals. Despite these problems, Pareto did not consider psychological knowledge as something unworthy to study. Quite the contrary – for him economic phenomena have a psychological explanation. However, economics in order to make progress must be a separate science and it needs to establish its fundamental laws (Bruni, & Sugden, 2007). Pareto thought that happiness and pleasure are too elusive terms and economics should not waste its time on the discussion about them. This escape from psychology and hedonistic foundation ended with Samuelson and his theory of revealed preferences.

Samuelson (1938) tried to base consumer choice theory on strictly observable foundations. However, the purpose of Samuelson's approach was not to 'reveal' preferences, but to create a strictly operational theory of consumer choice in which preferences or utility do not matter at all.³ He promised to remove the last "vestigial traces of the utility concept" (Samuelson, 1938, p. 61) from consumer choice theory. Samuelson was not interested in explanation (why people choose something), but he only wanted to describe the world. In order to do that we need to have a method of measurement. This is why Samuelson based his model on observable terms and consequences. He was operationalist, which means that his views were strongly influenced by logical positivism (Vienna circle) – the philosophical theory which had a considerable impact on value-free economics. Samuelson did not believe that the purpose of science should be finding the Truth. For him science had a more practical purpose. He wrote that "those who can, do science; those who can't prattle about its methodology" (Samuelson, 1992, p. 240).⁴

³ He writes that "The view that consumers maximize utility is not merely a law of economics, it is a law of logic itself" (as cited in Ainslie, 2001, p. 8).

⁴ Of course, Samuelson was not implying that methodology was unnecessary for economics, but he stated that if science worked, then no methodology would be required. Here this quote will not be discussed further because later the methodological discussion about this issue will be examined. Now it can only be stated that economist must deal with philosophy no matter whether they want this or not. Back in 1982 there was a debate between Tobin and Nozick, which illustrates this point. Tobin said to Nozick: "There's nothing more dangerous than a philosopher who's learned a little bit of economics." To which Nozick immediately responded: "Unless it's an economist who hasn't learned any philosophy" (Hutchison, 1996, p. 187).

The primary purpose of revealed preferences was to remove operational terms from economic theory (Hands, 2014). We need to have observable and measurable data to build a theory. It is impossible to observe preferences. Thus, we can only look at the act of choice, and on this assumption Samuelson based his theory. The basic idea of revealed preference is as follows. If one chooses option x instead of option y , then option x is revealed to be preferred to option y . Choices are consistent if they satisfy the ‘weak axiom of revealed preference’ (WARP), which means that, if x is revealed to be preferred to y , then y must not be revealed to be preferred to x . If choices satisfy the consistency conditions, then we can construct a complete, transitive and continuous revealed-preference ordering from them (Sen, 1971, 1973).

From the perspective of this book, the most crucial feature of revealed preferences theory is the perception of rationality. In this theory agents are rational when they act by the weak axiom of revealed preferences (WARP). Therefore, rationality is treated instrumentally, because economists do not know and do not want to know about the motivations of people. Neoclassical economists are only interested in the results but not in the causes of behaviour. The economists stopped being interested in human motives and they assumed *a priori* that people are always rational and that they always maximize their utility. However, it is essential to understand that utility here is entirely different from Bentham’s utility and utility used by marginalists. Nowadays, when economists say that individuals maximize utility, they are only saying that people do not rank any feasible option above the option they choose. Sometimes economists speak about individuals who are aiming to maximize utility or seeking more utility. However, it does not mean that utility is an object of choice or it is perceived as an ultimately (or absolutely) good thing. Utility is only an index of preferences. Utility maximizer only does what she prefers. To say that agents maximize their utility say nothing about the nature of their preferences. The only thing it does is to connect preference and choice. Rational individuals rank available alternatives and choose what they prefer the most (Hausman, 1992, p. 18).

In the end, however, it must be stated that Samuelson has not achieved what he wanted. In 1938 he tried to base consumer choice theory on strictly observable foundations which would differentiate his theory from both cardinal and ordinal approaches. In the paper from 1938, he did not introduce the term ‘revealed preferences’ (Samuelson, 1938). He started to use this term in 1948 (Samuelson, 1948). This change is important, because if well-ordered preferences exist, then they can be represented by an ordinal utility function. It means that utility is still in the interest of economics (Hands, 2009). In 1938 Samuelson perceived revealed preference theory as an alternative to utility theory. In 1950, however, it was obvious that this theory was nothing different from the ordinal approach. As Wong (2006, pp. 73, 74) summed it up:

The upshot of our interpretation of the purpose of the theory is that the revolutionary significance of the Samuelson Theory is lost. The development of the theory does not represent a break with the tradition in economic theory in which consumer behaviour is explained in terms of preferences (and material circumstances). Consequently, the attendant philosophical and psychological controversies of utility theory, which Samuelson hoped to evade with his observational theory, are not exorcised from the corpus of economic theory.

Although Samuelson was not able to introduce a theory different than ordinal utility approach, his efforts had an enormous influence on neoclassical economics which still perceives rationality in an instrumental way, has no interest in human motivations, and assumes that choice “reveals” human preferences.

1.3. Robbins and value-free economics

It is hard to overestimate the importance of Lionel Robbins' *Essay on the nature and significance of economic science* (Robbins, 1932). It was described as a “classic tract” (Caldwell, 1982, p. 99) and “the ‘official’ statement of the general ontology and epistemology of the discipline” (Ross, 2005, p. 87). Witztum (2007, p. 58) indicates three central tenets of Robbins' most important essay:

1. The subject matter of economics is the tension between scarcity and wants (the means-ends problem).
2. Economics is based on axioms (abstractions) which are derived from experience and which lead to statements about reality (hence, the ‘scientific’ nature of the subject).
3. Economics is not concerned with ends but only with the means available to achieve those ends. It is, thus, value free.

These tenets were necessary to achieve three goals. Firstly, they define economic science as the study of scarcity-constrained rational choice. Secondly, they put economics on a firmer epistemological foundation (moving away from hedonism). Thirdly, they make an argument against interpersonal utility comparisons (Hands, 2007). If these three goals had been achieved, marginalist economic would have been provided with exclusive rights to the title of scientific economics. Of course, the profession would still need historical data and other empirical evidence provided by the institutionalist or historical schools.⁵ However, it leaves the pure science of economics to marginalism alone (Hands, 2007). In this chapter, each of these goals will be analysed, especially in the context of value-free economics.

When it comes to the first assumption, Robbins writes that “economics is the science which studies human behaviour as a relationship between ends and

⁵ Robbins mentions “a Schmoller, a Veblen, or a Hamilton” (Robbins, 1932, p. 116).

scarce means which have alternative uses“ (Robbins, 1932, p. 16). This is probably one of the most recognizable quotations in economics. In these few words, it is explained what economics is all about. According to Robbins, man:

has to choose. He has to economise. Whether he chooses with deliberation or not, his behaviour has the form of choice. The disposition of his time and his resources has a relationship to his system of wants. It has an economic aspect. (Robbins, 1932, p. 12)

It goes without saying that time is limited and we cannot do a few things at the same time. Therefore, we always need to choose. This means that economics can deal with almost all issues (economic imperialism), and later Becker stuck fundamentally to this perception of economics. Before Robbins economists had a completely different understanding of economics. For them (e.g., Mill, Marshall, Smith) economics was a study of wealth. Robbins who was influenced by logical positivism could not agree with this. Firstly, the previous definition becomes a kind of tautology, because economists decide what wealth is. Therefore, Robbins underlines that this definition is incapable of describing its ultimate subject-matter (Robbins, 1932, p. 4). Secondly, the division between economic and noneconomic spheres is arbitrary, because if the conception of wealth is taken literally, the important parts of the generalizations of the science fall outside the area of economics (Robbins, 1932, pp. 5–8). This problem is visible on the example of Marshall and his perception of economics as the science of material welfare ((Marshall, 1920); see also (Dzionek-Kozłowska, 2007)). Robbins argues that many things are material without being economic (e.g., a pile of wood in the forest, worthless rocks, sand on the beach) and many other things are economic without being material (e.g., services). Due to these problems, it is understandable that Robbins’ definition is more precise.

Now, it is possible to analyse the most critical assumption from the perspective of this chapter – value-freeness. Before going deeper into Robbins’ ideas it is essential to understand how logical positivism classifies sentences. Therefore, it will be possible to comprehend the distinction between normative and positive approaches that still exists in economics. The logical positivists introduced a tripartite classification of sentences: 1) synthetic, 2) analytic, 3) nonsense (after Wittgenstein) (Putnam, 2002, pp. 10, 18). Synthetic judgments are empirically verifiable or falsifiable. They contain statements that can be verified by reality (facts). For example, we can say that Mount Everest is the highest mountain in the world. It is easy to check whether if this statement is true or false. We know how to measure height and we can measure all mountains in the world. The second type of statements are those which cannot be logically verified by reality. These analytical statements take the form of tautologies. This means that it is possible to indicate whether they are true or false on the basis of the logical rules alone. The purpose of tautologies is not to tell something about reality but

they show the logical structure of the world (Sady, 2013, p. 11). For example, we assume that the sum of angles in a triangle is 180 degrees or we can state that all bachelors are unmarried men. The third group of statements include all ethical, metaphysical and aesthetic judgments that are “cognitively meaningless” (Putnam, 2002, p. 10). Logical positivists perceive them as unscientific, because they are unverifiable. It is worth noticing that “This terminology [nonsense]⁶ is to be understood as implying a logical, not say a psychological distinction; its use is intended to assert only that the statements in question do not possess a certain logical characteristic common to all proper scientific statements” [namely verifiability]⁷ (Carnap, 1934, pp. 26, 27). Some philosophical problems arise in the case of ethics in which norms are used in the form of constative sentences, e.g., “killing is wrong.” People very often thought that this statement must be right or wrong. However, logical analysis proves that from the statement “killing is wrong” we cannot derive any statements which will tell about future experiences. Therefore, this statement is unverifiable and nonsensical (Carnap, 1935, § 1.4). It is important to bear this in mind, because logical positivists considered every value statements as meaningless. This reluctance towards value statements will be visible in the context of normative-positive dichotomy.

The distinction between positive and normative economics was popularized by John Neville Keynes. The separation of the positive and the normative was the part of Keynes’ more general distinction between positive science, normative science, and art. Positive science contains a study of fact (what is), normative science contains the study of norms and rules (what ought to be), and art is focused on policy application (what can be achieved). Keynes (1999, p. 22) writes about these differences:

As the terms are here used, a positive science may be defined as a body of systematized knowledge concerning what is; a *normative* or *regulative* science as a body of systematized knowledge relating to criteria of what ought to be, and concerned therefore with the ideal as distinguished from the actual; and art as a system of rules for the attainment of a given end. The object of a positive science is the establishment of uniformities, of a normative science the determination of ideals, of an art the formulation of *precepts*.

At the beginning of neoclassical economics, the difference between positive and normative approach was sharp, but economists did not undermine the necessity of a normative approach. They thought that ethics is necessary, but it should not be the area of interest for economics. One of the reasons for this belief was specialization. Economists could not deal with every problem, so they needed to specialize. This kind of reasoning was initiated by Jevons who wrote that “Various social sciences, also, are needed to promote the welfare of

⁶ Added by the author.

⁷ Added by Putnam (2002, p. 18).

mankind [...]. But Political Economy is distinct from all these other sciences, and treats of wealth itself" (Jevons, 1911, p. 7). Political economy takes wealth as the object of study, which is a component not the whole, of human welfare. Against those who complained about the narrow focus on wealth, Jevons (1911, p. 8) referred to the principle of division of labour:

But these complainers misunderstand the purposes of a science like political economy. They do not see that in learning we must do one thing at a time. We cannot learn the social sciences all at the same time [...]. There must be many physical sciences, and there must also be many social sciences, and each of these sciences must treat of its own proper subject, and not of things in general.

Until the 1920s, most economists believed that the relationship between economics and ethics was hierarchical. Economics was the science of wealth, and ethics used the insights from economics and the other social sciences to make judgments about which courses of action were ethically desirable (Yuengert, 2000). In value-free approach, economists are like engineers who tell individuals how to resolve some problems efficiently but they will not tell individuals what should they do.⁸ Economists only help people with achieving the goals that people set themselves. Economists are proud of this separation from ethics, because they can be perceived as objective scientists who deal only with facts. The argument of specialization was also used by Robbins who responded to Mill's assertion that good economics contributes to the ethical debate in these words "we may not agree with J.S. Mill that 'a man is not likely to be a good economist if he is nothing else.' But we may at least agree that he may not be as useful as he otherwise might be" (Robbins, 1932, p. 150). Specialization was an important factor in dismissing ethics from economics; but it was Robbins who established the dichotomy between normative and positive science. Moreover, he considered normative approach as unscientific and he believed that economists should not take ethics into consideration. Nowadays, the positive-normative dichotomy is widely recognized in mainstream economics. For example, the most famous introductory economics textbook of all time explains the difference between positive economics and normative economics in this way:

In thinking about economic questions, we must distinguish questions of fact from questions of fairness [...].

Positive economics deals with questions such as: Why do doctors earn more than janitors? Does free trade raise or lower the wage of most Americans [...].

Normative Economics involves ethical precepts and norms of fairness. Should poor people be required to work if they are to get government assistance? (Samuelson, & Nordhaus, 2001, pp. 7, 8)

⁸ Now we look at economists more as the plumbers who lay the pipes and fix the leaks. The conclusion from this metaphor is that economists should be more practical and check what works rather than creating grand theories that do not work in reality (see (Duflo, 2017)).

Hands (2012, p. 4) excellently explains how differently economists perceive these two approaches:

The dichotomy – the strict separation of the positive and the normative – was replaced by an epistemic condemnation and prohibition of the normative; not only was it necessary to recognize that positive and normative statements were fundamentally different, in addition it was argued that the normative was scientifically illegitimate and should be prohibited from proper economic science. According to logical positivism there were only two types of meaningful discourse – empirical science (synthetic knowledge) and logic/mathematics (analytic knowledge) – everything else was meaningless metaphysics. Since normative economics was based on presuppositions that were not derived from either of these two sources, normative economic science ceased to be any type of science at all, and was relegated to the epistemic dustbin along with religion, metaphysics, and other “meaningless” discourse. This positivist view of the normative was often combined with an emotivistic view of ethics – that ethical statements were simply expression of attitude or emotion.

Undoubtedly Robbins was influenced by logical positivism which assumed that ethical statements do not express propositions but emotional attitudes. Moreover, this view simplified all different value judgments to ethical ones. Due to logical positivism, debates about ethics were declared unresolvable (ethical values are simple statements of preference and, as such, are not the subjects of rational dispute) which meant for Robbins that economics had to be separated from ethics. He expressed the difference between ethics and positive economics in a few famous quotations. He writes:

Economics deals with ascertainable facts; ethics with valuations and obligations. The two fields of inquiry are not on the same plane of discourse. Between the generalisations of positive and normative studies there is a logical gulf fixed which no ingenuity can disguise and no juxtaposition in space or time bridge over. (Robbins, 1932, p. 132)

And he continues:

If we disagree about ends it is a case of thy blood or mine—or live or let live according to the importance of the difference, or the relative strength of our opponents. But if we disagree about means, then a scientific analysis can often help us resolve our differences. If we disagree about the morality of the taking of interest (and we understand what we are talking about), then there is no room for argument. (Robbins, 1932, p. 134)

Due to logical positivism, all value judgments were put into the one bag of nonsenses about which we cannot rationally discuss. Because of these problems, Robbins wanted to distinct economics from ethics and he argued that “economic analysis is *wertfrei* (value-free)” (Robbins, 1932, p. 91). Later economics became, in the words of Amartya Sen, “self-consciously non-ethical” (Sen, 1987, p. 2). Many years after the publication of Robbins’ essay, Kenneth Boulding, in his

presidential address to the American Economic Association, regretted that economics had evolved into a social science which is very different from moral science. He wrote:

We are strongly imbued today with the view that science should be *wertfrei* and we believe that science has achieved its triumph precisely because it has escaped the swaddling clothes of moral judgment and has only been able to take off into the vast universe of the 'is' by escaping from the treacherous launching pad of the 'ought'. (Boulding, 1969, p. 1)

It was almost 50 years ago and still a lot of economists preach false normative-positive dichotomy. Steven Levitt, the author of bestseller *Freakonomics*, wrote that "morality, it could be argued, represents the way that people would like the world to work – whereas economics represents how it actually does work" (Levitt, & Dubner, 2005, p. 13).

The next assumption of value-free economics is that we cannot compare utility between agents. Robbins used the prohibition of normative approach in his argument against utility comparisons. He wrote:

It is a comparison which necessarily falls outside the scope of any positive science. To state that A's preference stands above B's in order of importance is entirely different from stating that A prefers *n* to *m* and B prefers *n* and *m* in a different order. It involves an element of conventional valuation. Hence it is essentially normative. It has no place in pure science. (Robbins, 1932, p. 139)

It goes without saying that Robbins was a supporter of ordinal not cardinal approach towards utility. He believes that people can rank their preferences. How did he know that? Robbins answered that it was enough to look at everyday inner experience (i.e., introspection):

The main postulate of the theory of value is the fact that individuals can arrange their preferences in an order, and in fact do so [...]. We do not need controlled experiments to establish their validity: they are so much the stuff of our everyday experience that they have only to be stated to be recognized as obvious. (Robbins, 1932, pp. 78, 79)

For Robbins thanks to experience which is observable, ordinal approach has a solid scientific base in contrast to cardinal approach. Robbins' arguments support the revealed preference theory and they persuade many economists that value-free economics is an objective science.

1.4. Positive versus normative approach

In this section, it will be argued that a gulf between normative and positive judgments is not as wide as it is commonly thought by many economists (Blaug, 1992; Colander, & Su, 2015; Fiedor, 2013; Hands, 2012; Hardt, 2020; Kwarciniński, 2016; Mongin, 2006; Putnam, 2002; Putnam, & Walsh, 2011). In order to show

that, the misinterpretation of value judgments by economists will be presented. Moreover, it will be argued that normative assumptions and values exist in positive economics. However, many economists often believe in some normative axioms without even realizing that they are normative.

Nowadays, the dichotomy between positive and normative, facts and values seems straightforward, but for a long time it has been the subject of philosophical debate. Much of the discussion can be traced to David Hume and this is why the difference between normative and positive is also called 'Hume's dichotomy,' 'Hume's fork,' and 'Hume's guillotine.' Hume's primary concern was to block efforts to ground ethics in the facts of nature (Putnam, 2002, p. 2). In the famous passage Hume (1896, p. 319) writes that:

In every system of morality, which I have hitherto met with, I have always remark'd that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, is, and, is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is, however, of the last consequence. For as ought, or ought not, expresses some new relation of affirmation, 'tis necessary that it should be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.

Hume was right that there is no relationship between the statement 'I give to charity' and the statement 'I ought to give to charity.' That something *is* does not imply that it *should* be that way, and that it should be that way does not imply that it is. Hence, the difference between 'is' and 'ought' seems significant enough to call it a dichotomy (Hands, 2012).

However, the fact/value dichotomy comes out as something different than in Hume's writing. He did not use a term which we now call 'value judgments' (though he had the term 'morals'). Admittedly, he discussed individual value terms, for example, 'virtue,' 'ought' and so on. However, he put them into the context of ethics (Putnam, 2002, pp. 18, 19). Thus, he wrote about ethical judgments, not about all value judgments. Later a lot of philosophers forgot about this distinction and ethical statements started to be identified with value statements. This is visible in the case of the positivists. Carnap generally spoke not of "value judgments" but only of the statements of "regulative ethics" (or sometimes "normative ethics") (Putnam, 2002, p. 18). The goal of positivists was to challenge objectivity and rationality of ethics. So, their fact/value dichotomy was not based on any serious examination of the nature of values or valuation at all. They examined the nature of "fact" in a narrow empiricist spirit (Putnam, 2002, pp. 28, 29). Therefore, the proper analysis is necessary if we want to understand the nature of value judgments.

Hands indicates that normative statements are something more than ethical valuations. He writes:

In general normative terms are simply terms that are action-guiding or prescriptive, and normative statements are statements involving such terms. The relevant norms might be social, legal, epistemological, aesthetic, or a host of other types; ethical norms are just one very special case of such prescriptive terms. (Hands, 2012, p. 9)

Outside economics, it is obvious that not everything that is 'normative' involves ethics. For example, if someone says 'you ought to eat more vegetables' it does not mean that you should eat them to be more moral. It means that you ought to eat vegetables to be healthy. Therefore, it is a norm but the norm of good health, not an ethical norm. Economists have problems with this distinction, which is visible in the case of welfare economics where any value component is automatically considered as ethical. For Mongin the problem is deeper than it looks at first glance. He writes:

The economists' confounding of ethical with non-ethical evaluations is perhaps only secondary; they are trapped into confusion at an earlier stage. They think of evaluations in terms of judgments, not propositions. This psychological preconception fosters the belief that there cannot be an evaluation without a person's commitment, and personal commitments are supposedly typical of morality. (Mongin 2001, p. 13)

Therefore, if one says "I do believe that recent actions of government were positive" it does not always mean that one wants to evaluate their actions but it could also mean that one gives information to someone about the actions of the government.⁹

The other side of the problem is that many economists use normative assumptions without realizing that. However, before going to explain what it means we need to understand that economists sometimes unknowingly derive 'ought' from 'is.' It is possible, because economics has been significantly influenced by utilitarianism which is a consequentialist system. For a utilitarian one ought to do what gives one the most pleasure. Therefore, if *A* causes more pleasure than *B*, then *A* ought to be done. In this way, utilitarianism connects economic evaluations with prescriptions (Mongin, 2006). Of course, these assumptions lead economists to choose a particular policy. Hands shows how this process works. If an economist had to choose between policy *A* and *B* he would probably use this criterion "The best (good) policy (*ceteris paribus*) is the one that makes people better off (that is, it leads to the highest level of utility or preference satisfaction

⁹ Terms like 'honest,' 'generous,' or 'courageous' can be perceived as thick predicates of morality. The defining feature of thick predicates is that they have both an evaluative and a descriptive side in an opposition to ethical good which is typically thin predicate (Mongin, 2006).

among the relevant individuals)" (Hands, 2012, p. 16). So, if economic analysis tells us that policy *A* (*ceteris paribus*) makes people better off than policy *B*, then society ought to do *A*. Of course, economics does not indicate what things make people better off. However, it assumes that what people do maximizes their utility. In the end, it means that economics knowing what is (what people do – positive facts) tells us how it ought to be (normative).

The next problem also arises from is/ought dichotomy. For economists this distinction proves that economics is positive and value-free. Klappholz (1964, pp. 98, 99) points that out:

The 'orthodox' position rests on Hume's observation that norms or proposals cannot be deduced from descriptive statements alone [...] Since 'the scientific part of economics consists exclusively of descriptive statements, it cannot have any ethical entailment, and is therefore value-free'. If this is what we mean by 'value-free' then, as J. N. Keynes said, it seems trivially true that economics, and the social sciences in general, can be value-free.

Moreover, Mongin highlights that "positive economics starts from assumptions that are phrased in terms of factual statements alone and it proceeds by making deductions from these statements" (Mongin 2001, p. 20). However, this understanding is flawed, because Hume's dichotomy only works when initial premises are exclusively factual and economics contains the evaluative vocabulary of preference. The statements like 'I prefer *a* to *b*' designate an evaluative concept which makes the proposition an evaluative one (Mongin, 2001).

Having explained these problems, it is now possible to scrutinize two concepts which are essential for neoclassical economics: Pareto optimum and rational choice theory. Pareto optimum shows how neoclassical economics perceives welfare. Firstly, this concept is based on methodological individualism. Look at these two statements:

- (1) state *x* is better than *y* for individual *i*,
- (2) state *x* is better than *y* for society.

In the first case, betterness is connected with the preferences of people. Therefore, the state which satisfies people's preferences more is the better one. The second example is problematic, because we do not know what satisfies preferences of society. We can only refer to individual preferences and then sum them up (Mongin, 2001). However, from the 1930s and 1940s, economics ceased to make interpersonal utility comparisons and started to use the Pareto criterion. According to this theory, a specific distribution of resources is Pareto Optimal (efficient) only when reallocation that would make one person better off, would also make someone else worse off. This idea was the basis for the 'First Fundamental Theorem of Welfare Economics' in which any competitive equilibrium leads to Pareto efficient allocation of resources. Why did economics

shift towards Pareto criterion in the 1930s? Because it helped economists to avoid the troublesome normative issues associated with interpersonal comparisons of utility. Thanks to that, economists look more 'scientific.' They were able to make positive/scientific judgments about social welfare and microeconomic policy (Hands, 2012). However, Pareto optimum (or economic efficiency) is normatively laden theory and welfare economics cannot be perceived as positive science (Blaug, 1992; Hausman, 1992; Hausman, & McPherson, 2006). Pareto criterion is not value-free, because it is based entirely on satisfying the preferences of the particular economic agent. Therefore, being better or worse off is caused by individual preference satisfaction. For Hausman and McPherson, Pareto criterion functions "as a Trojan horse smuggling ethical commitments into the theoretical citadel of positive mainstream economics" (Hausman, & McPherson, 2006, pp. 67, 68). Many economists believe that economic efficiency is value-free and it allows them to "conclude that, *ceteris paribus*, perfectly competitive equilibria are morally desirable and market imperfections that interfere with the achievement of competitive equilibria are morally undesirable" (Hausman, & McPherson, 2006, p. 66). This view has a significant influence on a contemporary free market society.

The second issue is rational choice theory and the fact that many economists consider it a positive approach. It would be a descriptive theory if economists made a systematic empirical observation of agent's preferences and choices, then generalized those observations to derive some scientific laws of economic behaviour (Hands, 2012). However, this is not how economic analysis is done in most cases. As it was mentioned before, economists cannot observe people's preferences, beliefs, and desires. Therefore, they are concentrated on people's choices. However, neoclassical economists not only state that someone chooses something but they also assume that someone chooses something, because it maximizes one's utility or satisfies one's preferences. Moreover, neoclassical economists assume that one's choice is rational. Economists are satisfied to use instrumental rationality which enables them to be value-free, because they do not state which choice people should make to maximize their utility. From that perspective, economics only deals with means not ends that people put on themselves. However, some economists forget that they make normative assumptions about instrumental rationality. They assume that preferences must be transitive and complete to be called rational. These assumptions are treated as positive by many neoclassical economists and this is the reason why Thaler's article *Toward a positive theory of consumer choice* surprised so many of them. Thaler (1980, p. 39) points out that:

Economists rarely draw the distinction between normative models of consumer choice and descriptive or positive models. Although the theory is normatively based (it describes what rational consumers should do) economists argue that it also serves well as a descriptive theory (it predicts

what consumers in fact do). This paper argues that exclusive reliance on the normative theory leads economists to make systematic, predictable errors in describing and forecasting consumer choices.

One of the goals of behavioural economics was to demonstrate the situations when people behave differently than the neoclassical theory assumes.¹⁰ Although the criticism of neoclassical economics is severe, later it will be examined how economics thanks to Becker was able to defend their hard core assumptions (maximization of utility, instrumental rationality) and absorb descriptive criticism. Even though behavioural economics show that rational choice theory does not represent how people behave, many economists consider it to be positive, not a normative theory. It seems even stranger when we take into consideration that philosophers perceive economic rationality as a normative theory (Sugden, 1991). However, it is not strange at all if we realize that over the last dozen years economists and philosophers have not had a lot of contact with each other. However, this is changing now. Due to the philosophy of economics, normative characterization of rationality has been more recognizable among economists (e.g., (Blaug, 1992; Boland, 1997; Davis, 2003, 2010; Hands, 2012; Hausman, 1992; Hausman et al., 2017; Mongin, 2006; Reiss, 2013; Ross, 2005)).

1.5. Conclusions

To summarize this chapter, it can be stated that the distinction between *is* and *ought* is exaggerated in economics and too strong conclusions are derived from it. Following the advice of the philosopher Hilary Putnam, the dichotomy should be ‘disinflated:’

If we deflate the fact/value dichotomy, what we get is this: there is a distinction to be drawn (one that is useful in some contexts) between ethical judgments and other sorts of judgments. This is undoubtedly the case, just as it is undoubtedly the case that there is a distinction to be drawn (and one that is useful in some contexts) between chemical judgments and judgments that do not belong to the field of chemistry. But nothing metaphysical follows from the existence of a fact/value distinction in this (modest) sense. (Putnam, 2002, p. 19)

Moreover, economists must understand that they are not objective and positive scientists even though many of them see themselves as such. Economic theories (rational choice theory, Pareto optimum) contain normative components but to realize that economists must extricate themselves from positivistic influence

¹⁰ For example, dictator game/ultimatum game (Kahneman, Knetsch, & Thaler, 1986), mental accounting (Thaler, 1985), prospect theory (Kahneman, & Tversky, 1979), hyperbolic discounting (Ainslie, 1975), inner motivations crowding-out (Frey, & Oberholzer-Gee, 1997). The most important achievements of behavioural economics are presented in (Thaler, 2015).

(e.g., value judgments are not ethical judgment). In general, it seems justified to point out that specialization has not been always profitable for economists. They need to listen to other social scientists and ethicists, because economists are not value-free engineers who deal only with means, not ends. No social science can be entirely value-free and scientists deal with human goals in the end. Therefore, it is better to do it properly rather than not do it at all. To do that properly, economists need to have an honest debate on goals and normative approaches. However, before we move to normative approaches and their methodological justification, we need to analyse Becker's economic approach, because this approach enables economists to rebut every kind of descriptive criticisms towards the 'hard cores' of value-free economics (maximization of utility, instrumental rationality). And so, economics still is perceived by many economists as a positive and value-free science.

Chapter 2

Lakatosian perspective on Becker's economic approach

*To abandon neoclassical theory
is to abandon economics as a science*
(North, 1978, p. 974)

2.1. Introduction

Until the 1970s methodological views among the majority of economists were based on the philosophy of logical positivism. Especially, they preached the hypothetico-deductive model of scientific explanation which was introduced by the Vienna circle (Blaug, 1992, p. 4; Caldwell, 1982, pp. 11–18). These ideas were brought to economics mainly by Hutchison (1938). After the Second World War positivism had the most influence on economics. Especially, after the publication of Friedman's seminal essay *The methodology of positive economics* (Friedman, 1953). However, over the last few decades, the post-positivists philosophers of science (e.g., Popper, Kuhn, Lakatos) had gained importance. From the perspective of our analysis, one of them (Lakatos) is essential.

Lakatos' main idea is that the unit of scientific achievements is not an isolated hypothesis but a 'scientific research programme.' Scientific research programme consists of (1) hard core, (2) protective belt, and (3/4) positive/negative heuristics. (1) Hard core contains the fundamental metaphysical assumptions of the research programme. It is a framework of general hypotheses which elements are irrefutable by the empirical evidence. Hard core remains fixed in spite of programme's development, because the rejection of hard core means the abandonment of the programme itself (Hands, 2001, p. 122). (2) Protective belt contains auxiliary hypotheses, empirical conventions, and other theoretical structures of the programme which may be falsified. In this area, changes occur as programme moves through time. Protective belt is a buffer zone between hard core and empirical evidence. Protective belt changes in response to changes in empirical evidence (Hands, 2001, p. 122). (3/4) Positive and negative heuristics provide information about what should (positive) and should not (negative) be sought during the development of the programme. They constitute a set of

accepted methodological rules (Lakatos, 1980, p. 47). More specifically, positive heuristic is a “set of suggestion or hints” which help with adjusting sections of the protective belt (Lakatos, 1980, p. 50). Negative heuristic veers criticism about falsity or inadequacy towards the protective belt which makes hard core irrefutable (Lakatos, 1980, pp. 48–50). Lakatos was influenced by Kuhn (1962) and his model is consistent in some areas with Kuhn’s view of science. Usually, most things that happen in scientific research programme take place in protective belt, whereas hard core remains untouched. A scientific revolution occurs when hard core is replaced. For Lakatos scientific progress takes place when degenerating research programme is replaced by new progressive one which leads to the discovery of novel phenomena (for more information see (Lakatos, 1980; Redman, 1993)).

At the beginning of the 1970, the idea of scientific research programme gained popularity among economists. It was possible, because Lakatos answered problems to which Popper (1959) did not. It was the case in the areas in which a tension between falsificationism and the actual practice of economics exists. This problem was captured by Blaug who wrote that “the central weakness of modern economics is in fact the reluctance to produce theories which yield unambiguously refutable implications” (Blaug, 1975, p. 425). It goes without saying that neoclassical economics contains many hard cores which define the research programmes used by economists (Hands, 1993). However, the conception of scientific research programme did not fit into economics as perfectly as it was thought at the beginning. The decline of Lakatos’ influence on economics started at Capri conference in 1989. Many economists interested in economic methodology (e.g., Backhouse, Blaug, Caldwell, Hands, Hoover, Mäki, de Marchi, Mirowski, Morgan, Smith, Weintraub) met there (see (Gonzalez, 2014)). The result of this conference was the book *Appraising economic theories* (de Marchi, & Blaug, 1991) which was mostly negative towards Lakatos. Since then, Lakatos’ method has lost its popularity among economic philosophers to the present day (see (Backhouse, 1994; Hands, 1993; Hausman, 1992, ch. 6; Rosenberg, 1992, ch. 4)). Why has it happened? The purpose of this section is not a general methodological discussion on science and its limitations. It is constrained to two reservations towards Lakatos which were summarized by Blaug. Firstly, it is impossible to establish the hard cores of any research programme in economics about which everyone would agree.¹ Secondly, measuring scientific progress by empirical criterion (prediction of novel facts) is inappropriate for economics (de Marchi, & Blaug 1991, p. 500). This argument was also underlined by Hands (1984, 1990) who argued that social science like economics lacks “crucial experiments” which would enable us to predict novel

¹ As an example see the critical analysis of “neo-Walrasian research programme” (Weintraub 1985, p. 109) by Hausman (1992, pp. 87, 88).

facts. This point is crucial, because for Lakatos prediction of novel facts gives scientists a chance to compare and decide which rival scientific programme is better (Hands, 1990). Although Lakatos' scientific research programme cannot indicate hard core and which programme is better with absolute certainty, it still can be a useful method. Despite the criticism presented above, many economists continue to employ Lakatosian models of methodological explanation in almost all fields of economics (Drakopoulos, & Karayiannis, 2005; Hands 2001, pp. 299, 300). Lakatos is still popular in economics, because his idea gives economists a chance to identify the structure of various research programmes. Identification of hard core, protective belt, positive heuristic, and negative heuristic helps organize our thoughts about numerous approaches in economics. Moreover, thinking in categories of progress or deterioration can help (to some degree) with assessing whether some research programme is useful or not.

This analysis can be especially fruitful in the case of Becker's economic approach. There are a few reasons why we should analyze the economic approach in the context of scientific research programme. For some economists, Becker represents accurately what is called 'neo-classical economics.' This identification is supported by Becker himself (1976) as well as other writers (Coleman, 1993, p. 169; Hands, 2001, p. 273). Others (e.g., (Blaug, 1992, p. 221)) refer to Becker as part of the Chicago school. Generally, Becker is considered to be connected with "the standard," "mainstream" or "neo-classical mainstream" of economics (Hands, 2001, pp. 271–273). Overall, he is regarded as the 'face' of neoclassical economics by other social scientists. It is difficult to overestimate the importance of Becker. According to Fuchs, Becker is "one of the most influential social scientists of the second half of the 20th century [...] [whose] influence on the other social sciences exceeds that of any economist of his generation" (Fuchs, 1994, pp. 183, 184).

Having this in mind, it is surprising that more attention was not paid to Becker in the books about economic methodology. In a widely cited book of Hands (2001), Becker appears only in an indirectly laudatory introductory quotation by Boland (1997, p. 1) and then in a critical case study in a section about feminist economics (Boland, 1997, pp. 270, 271). Even less information we can find in Hausman (1992) where Becker is only shortly mentioned a few times. We can find more about Becker in Blaug (1992). He writes about Becker's theoretical and methodological contributions to marginal productivity theory (p. 175), human capital theory (p. 207), methodological individualism (pp. 209–211), and the new economics of the family (pp. 220–228). Moreover, there is comprehensive literature that analyses economic imperialism (e.g., (Lazear, 2000; Mäki, 2009a; Mäki, Walsh, & Pinto, 2017)). However, Becker's approach has not been thoroughly investigated from the perspective of scientific research programmes.² It is quite startling

² The partial analysis in (Blaug, 1992).

that the economic approach was not eagerly examined in the context of Lakatos' research programme, because it can be easily done. The economic method has some axioms from which the meaningful predictions about human behaviour are deduced. Becker writes that "the combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it" (Becker, 1976, p. 5). He believes that "the economic approach provides a valuable unified framework for understanding all human behavior" (Becker, 1976, p. 14). However, understanding here does not mean common sense understanding. For Stigler and Becker (1977) economic method must be judged in terms of its predictive power, not in terms of the descriptive realism of its assumptions or explanations (see (Becker, 1993, pp. 402, 403)). It fits into scientific research programme framework in which hard core contains assumptions that do not need to reflect reality. For Becker understanding human behaviour means being able to make predictions, and in this regard we can observe the influence of Friedman's book to which we referred earlier in this section (Friedman, 1953).

The second person who had an enormous influence on Becker was Robbins. His impact is observable in two areas. Firstly, Becker believes that economics could and should be value-free. Therefore, economics should not assess people's goals and decide which things are worth pursuing. Secondly, Becker based his method on the notion of scarcity. For him people always have to choose, because they cannot do two things at the same time, so time is a scarce resource. Due to this understanding, economics deals with every human choice (religion, love, childbearing, etc.). This is why economics started to analyse problems of other social sciences. This process is explained by Radnitzky and Bernholz who write:

Acting rationally, using appropriate means to achieve one's ends, involves essentially taking into account 'costs' and 'benefits', whereby 'costs' and 'benefits' are taken in its widest sense. Since resources are always scarce - last not least our life-time - rational conduct is governed by principles of economics. Hence, it is plausible that new and important knowledge can be gained by applying the perspective of economics, the methods and tools of economics - some of them suitable generalized - to fields of enquiry that have traditionally been thought to lie outside the competence of economics. (Radnitzky, & Bernholz, 1987, p. VIII)

Precisely for imperialism Garry Becker was awarded Nobel Memorial Prize. The prize was awarded for "having extended the domain of economic theory to aspects of human behaviour which had previously been dealt with - if at all - by other social science disciplines such as sociology, demography, and criminology" (Royal Swedish Academy of Sciences, 1993, p. 1).

Becker with his economic approach had an enormous influence on other social sciences and the reason for that lies in hard core assumptions of this method.

These axioms and a belief in being positive science give economics the dominant role among the other social sciences. However, in this chapter, the author will check whether the economic approach is the dead end for economics. Although it provides an answer to every problem, these answers are not necessarily correct. Moreover, it is worthwhile to analyse the Becker's economic approach in the context of scientific research programme to investigate how economists reacted to the criticism from the other social sciences. Due to the protective belt, they were able to rebut every kind of descriptive criticism and hard cores remain unchanged. Nowadays, still many economists are reluctant towards normative approaches or methodological pluralism. This is the result of a coherent economic method that can 'explain' everything. Thus, the question if economics should stick to its hard cores is more than justified.

2.2. Becker's economic approach as scientific research programme

Lakatos' scientific research programme will be used as a framework to understand how Becker's economic approach works and how it was changed by the criticisms towards neoclassical economics. Thanks to the conclusions from this investigation, it will be possible to answer the question if this research programme should be replaced by something else. Becker's economic approach can be easily perceived as research programme. Becker indicates 'hard core' of his approach when he writes that "the combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it" (Becker, 1976, p. 5). However, in this section hard core of the economic approach is maximization of utility and instrumental rationality. This incongruity between Becker's hard cores and those indicated in this book is possible, because it is difficult to establish hard cores to which everyone would agree (de Marchi, & Blaug 1991). Of course, hard cores presented in this section (maximization of utility, instrumental rationality) are open for further discussion. Before analysing each of the hard cores a few words of explanation are necessary. First and foremost, in this section equilibrium is not considered as hard core, because our focus is on how economics perceives a human being. However, undoubtedly this assumption is often used as a justification for a free market system and it is important to keep in mind that effectiveness is not a positive statement but a normative one. The second assumption stable preferences will be analyzed as 'negative heuristic.' The third assumption maximizing behaviour is divided into maximization of utility and instrumental rationality even though Becker perceived them as the same thing. He writes that "by 'behaving rationally,' I mean 'maximizing' consistent behavior that looks forward and tries to anticipate as far as possible what the future will bring. This is common to all versions of rational

choice that I know of" (Becker as cited in Swedberg, 1990, p. 40). The distinction between instrumental rationality and maximization of utility is made, because it is important to point out that utilitarianism had a huge influence on economics. Moreover, the fact that economics perceives rationality in an instrumental way is something worth noticing, especially if we look at how rationality is perceived among the other social sciences. The purpose of the analysis of the economic approach in the context of scientific research programme is to show how thanks to the expansion of the protective belt, the economic approach could absorb every criticism of its hard cores.

(1) Maximization of utility. Becker did not hide how much he derived from utilitarianism and Bentham. For Nobel Prize winner it was also crucial to assume that individuals are always seeking to maximize their utility. Becker argues that "everyone recognizes that the economic approach assumes maximizing behavior more explicitly and extensively than any other discipline, be it the utility or wealth function of the household, firm, union, or government bureau that is maximized" (Becker, 1976, p. 5). According to Becker, more important than utility measurement is how utility is perceived. He understands utility *ad libitum*, which means that utility could mean everything. By broadening the conception of utility, Becker answered two different kinds of criticisms of *homo economicus* (Ostapiuk, 2017b). Firstly, he embraced all altruistic behaviours into the process of maximization. From its beginning, and especially after the marginalists revolution, economics had been accused of perceiving people as selfish creatures who do not think about the others. Becker did not identify maximization of utility with self-interest. He argued that "individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, spiteful, or masochistic" (Becker, 1993, p. 386). This relativistic perception of utility causes that this term started to be the 'black box' which could contain everything. Therefore, it does not matter if people are altruists or egoists, because in the end they always maximize their utility (Sen, 1977). Secondly, "the extended conception of utility causes that we cannot establish any objective values which human beings want to pursue. It is impossible to compare utility between people and also between any choices made by the particular person." (Ostapiuk, 2019b, p. 641). In this respect, it is visible that Becker strongly supported the ordinal approach and value-free economics. What is even more important than maximization of utility is a consequential way of thinking used in the economic approach. There are no goods or goals which have value *per se*. The only thing that matters is a choice and when people choose they maximize their utility. Perceiving utility *ad libitum* creates a protective belt. It is obvious that sometimes people make bad decisions and they do not maximize their utility (e.g., taking hard drugs, smoking, and driving after drinking). These bad choices are anomalies and they pose a threat to hard core assumptions that people always maximize their utility. Therefore, utility is perceived *ad libitum*. As a consequence, all behaviours can be

explained as actions which maximize utility. In this way, hard core is saved from anomalies, because it is impossible to behave in a way which does not maximize utility. The impossibility of falsifying the maximization of utility assumption was indicated many years ago (Boland, 1981; Hodgson, 2012b; Sen, 1977).

(2) Instrumental rationality. Mainly due to the 'formalist revolution' (see (Blaug, 2003)) neoclassical economics perceived people as hyper-rational agents that can compute an infinite amount of information and decide which behaviour will maximize their utility. However, perfect rationality was an object of criticism which has intensified since behavioural economics gained popularity in the 1970s and 1980s. Nowadays, neoclassical economists do not use perfect rationality but instrumental rationality. It is a less demanding version of rationality and it is used to protect hard cores of the economic approach. It is obvious that people do not always compute all the information and on this basis make the best choices. Often people behave in irrational way. Economists have two ways of dealing with these anomalies.

Firstly, economists believe that people would make perfectly rational choices if it were not for various constraints like limited time or limited cognitive powers. Sometimes, economists argue that full rationality can be irrational. It is the case when the costs of acquisition and processing of information are too high. Then, as Knight noticed, "it is evident that the rational thing to do is to be irrational, where deliberation and estimation cost more than they are worth" (Knight, 1921, p. 67). Hayek was the economist who was probably the most consistent in praising the gut feeling and hidden knowledge in the situations when someone made a prompt decision without hours of deliberation. He wrote that "if we stopped doing everything for which we do not know the reason or which we cannot provide a justification [...] we would probably soon be dead" (Hayek, 1988, p. 68). Nowadays, heuristics³ – the simple and efficient rules which people often use to form the judgments and to make decisions – are not considered to be irrational, because people do not have enough time to digest every available information. If we had to think about our every move, we would not be able to live, because every day we must make thousands of different decisions. This is one of the conclusions of Kahneman (2011). He proves, among other things, that fast 'system 1' is more prone to errors than deliberate and related to homo economicus 'system 2.' However, these two systems are interrelated and people cannot only rely on system 2, because most of the time they need to make fast decisions (Kahneman, 2011).

The most prominent scholar that challenged the neoclassical assumption of perfect rationality was Herbert A. Simon. He argued that people were satisficing, not maximizing "whereas economic man maximizes – selects the best alternative

³ The term 'heuristics' was popularized by Tversky and Kahneman (1974).

from among all those available to him, his cousin, administrative man, satisfices – looks for a course of action that is satisfactory or ‘good enough’” (Simon, 1957, p. XXV). People do not maximize their utility but make satisfactory decisions due to the bounded rationality. People have bounded rationality, because there is not enough time to process all important information and our cognitive power is limited. Nevertheless, Simon’s idea was only an attempt to adjust *homo economicus* to reality, not its rebuttal. Bounded rationality does not dismiss the hard core assumption (instrumental rationality), because in the economic approach rationality is bounded. In this approach people still want to maximize their utility. However, due to the complexity of the external world, they cannot always maximize it.

The second way of dealing with irrational behaviours is a change in the understanding of the very term rationality. In the common view, rationality means wise and thoughtful choice. However, neoclassical economics is value-free and does not discuss goals. It assumes that it is impossible to rationally decide what people should want. This is why economists can only check if people satisfy their preferences effectively. Hume provides very famous description of an instrumental rationality. He writes that “reason alone can never be a motive to any action of the will;” and “reason is, and ought only to be the slave of the passions” (Hume, 1896, pp. 282, 283). Value-free economists treat these words very seriously and rationality in economics is not about outputs but the process. It does not matter what people choose, but if the process of choice meets two conditions: transitivity and completeness.

In the economic approach two lines of reasoning which explain anomalies are used: (a) full rationality is impossible, because of limited time and limited cognitive powers, (b) rationality is perceived as the process not output). However, especially the first one is crucial for Becker. He writes that “actions are constrained by income, time, imperfect memory and calculating capacities, and other limited resources, and also by the opportunities available” (Becker, 1993, p. 386). These limitations (especially limited time) are negative heuristics which create a protective belt. Stupid and irrational decisions do not challenge hard core assumption of rationality, because rationality in value-free economics is always limited.

2.3. Becker's methodology

Although Becker did not have a consistent methodological position, it can be useful to try to understand his methodological views to see who influenced him and how. Thanks to this analysis, we can understand why the economic approach works as it works. Sometimes Becker is perceived as positivist and this is the result of his value-free approach. However, he did not assume that every

economic theory has to be falsifiable or testable. What is testable is a group of assumptions when they are applied to particular problems. Then, we can assess if our prediction is accurate or not. Predictive power is essential in the economic approach and this is why Becker can be perceived as an instrumentalist. He was profoundly influenced by Friedman who is perceived as methodological instrumentalist (see (Boland, 1979; Caldwell, 1980, 1982)).⁴ Becker dedicates his most important book to Friedman from whom he 'learned the economic approach' and he expressed his gratitude for his comments (Becker, 1976, p. 3).

Instrumentalists do not care about unrealistic assumptions. For them, hard core assumptions are taken on faith but faith is based on functionality. What is more, frequently realistic assumptions are not good for the predictive power of a given theory. This view is identified with Friedman, who wrote that "in general, the more significant the theory, the more unrealistic the assumptions" (1953, p. 14). Friedman shows how he perceived the explanatory role of scientific theories when he writes:

A hypothesis is important if it 'explains' much by little, that is, if it abstracts the common and crucial elements from the mass of complex and detailed circumstances surrounding the phenomenon to be explained and permits valid predictions on the basis of them alone. To be important, therefore, a hypothesis must be descriptively false in its assumptions; it takes account of, and accounts for, none of the many other attendant circumstances, since its very success shows them to be irrelevant for the phenomena to be explained. (Friedman, 1953, p. 14, 15)

For Friedman it is not important to understand all phenomena in a theory. He writes that "A meaningful scientific hypothesis or theory typically asserts that certain forces are, and other forces are not, important in understanding a particular class of phenomena" (Friedman, 1953, p. 40). To successfully predict human behaviour it is enough to assume that people behave as if. This idea was presented in the famous example of a billiard player (Friedman, 1953, p. 21). Becker uses this concept eagerly when the assumption that "participants [...] maximize their utility" is transformed into the assumption that economic agents act 'as if' they maximize their utility (Becker, 1976, p. 14).

Despite the appraisal of the economic approach as a "foundation for predicting" (Becker, 1976, p. 6), Becker also states that "the economic approach has numerous implications about behavior that could be falsified" (1976, p. 10). It means that Becker accepts the importance of falsification. His writings contain normative criteria of Popper's doctrine. Becker writes approvingly of audacity, specificity, and falsifiability of predictions. Moreover, he denounces sweeping generality,

⁴ However, Mäki argues that Friedman's position is more complicated (2009b), see also (Hoyningen-Huene, 2017).

ad hoc hypotheses, and unwillingness to specify grounds for refutation (Turner, 2010). Falsificationism is consistent with instrumentalism, but it is possible, because instrumentalism does not provide any explicit method to assess the “testability” or “usefulness” of assumptions (Turner, 2010, p. 59). This is why Becker is more consistent with Lakatos who approves a “whiff of ‘inductivism’” (Lakatos, 1980 p. 159). However, Becker tried to adjust his economic approach to falsificationism and he even stated that the “implications” of the economic approach “all appear to be consistent with the available evidence” (Becker, 1976, p. 10). This need for testability of his theories will be used to show the internal inconsistency of Becker's approach when it comes to choices over time.

In addition to instrumentalism, Friedman taught Becker how comprehensive and powerful economics could be in comparison to the other social sciences. Becker writes “he [Friedman – A.O.] made me see that you can attack social problems with economics” (Becker as cited in Swedberg, 1990, p. 29). It was possible, because the objective of economics theory is to “‘explain’ much by little” (Friedman 1953, p. 14). Becker very often pointed out that the economic approach was “comprehensive” (Becker, 1976, p. 8) and a “unified framework” (Becker, 1976, p. 14) which enables “understanding all human behavior” (Becker, 1976, p. 14). This bold assertion can be understood as the extension of the search for “novel facts” (Lakatos, 1980 p. 5). However, it is closer to be an attempt to even larger objective in science which is a unification of explanation (Mäki 2009a; Sen as cited in Swedberg, 1990, p. 264). Power to predict new facts comes from simplicity (minimum of assumptions) and fecundity (precision, magnitude) (Turner, 2010, p. 70). For Becker and value-free economics, in general, these assumptions are achieved by rational choice theory. Becker writes that “the rational choice model provides the most promising basis presently available for a unified approach to the analysis of the social world by scholars from different social sciences” (Becker, 1993, p. 403).

The economics' primacy and its difference from the other social sciences come from the way in which economics looks at reality. Becker asserts that “what most distinguishes economics as a discipline from other disciplines in the social sciences is not its subject matter but its approach” (Becker, 1976, p. 5). This approach is simple and unified, because it does not need to deal with normative statements and have hard core assumptions. We can observe the ‘superiority’ of economic approach on the example of crime and punishment. Coleman points out that sociologists and criminologists analyse these issues from the normative preservice, whereas Becker analyses them from the “perspective of positive social science” (Coleman, 1993, p. 172). The normative view was that “no offenders should in principle be permitted, and all offenders should in principle be punished” (Coleman, 1993, p. 172). Becker transformed this into the positive question: “how many offenses should be permitted and how many offenses

should go unpunished?" (Becker, 1976, p. 40). Therefore, Becker discusses only positive notions such as efficiency, scarcity of resources in law enforcement, and minimization of the social losses resulting from crime. Whereas the dominant sociologists, as Coleman highlights, "saw crime as the simple product of external causes, the criminal as passive clay, and the goal of criminological policy as rehabilitation," Beckerians have used a "rational choice orientation" to show "the deterrent effect of punishment and [...] the effect of incentives on crime" (Coleman, 1993, p. 172).

2.4. Problems with instrumental rationality and utility maximization

Due to the numerous anomalies, the protective belt had to be expanded to defend the hard cores of the economic approach. Although the protective belt makes the economic approach work, it causes substantial problems. First of them is connected with utility maximization. Value-free economics does not deal with human goals and it perceives utility *ad libitum*. This relativistic approach results in a situation where economics becomes insensitive to the difference in human motivations.

It is essential here to underline that preferences go beyond utilitarian measurement. Arrow defines them as "the values of individuals rather than to their tastes" (Arrow, 1951, p. 23). Values are the core of human preferences and they form people's goals and preferences especially in the long run. Hechter describes them as "relatively general and durable internal criteria for evaluation" (Hechter, 1992, p. 215). Value-free economics wanted to save its utility maximization theory by putting all values into the 'black box' of utility. However, this way of thinking has led to tautology and it has made economics insensitive to differences in human motivations. Values are different from utility, because they are *ex-ante* reasons why people do something and they do not do something else. In this instance, we do not use consequential thinking, and we do not calculate our utility. Solving a chicken or egg dilemma is impossible. It means that we cannot prove that humans are solely motivated by values and that they are not influenced by utility. However, there is no need to prove that values are something that exists *ex-ante*. The author only wants to highlight that values vary from utility and we cannot compare them within the same framework. This difference is visible if we analyze people's attitudes towards the concepts of values and utility. One can use a value-free line of argument that people give back a wallet which they had found on the street, because it is beneficial to them to give it back. Cherishing one's self-respect, being a 'good person' can be more important than money in this economic way of thinking, because it gives one's more utility. However, this reasoning looks only as an

attempt to defend the *homo economicus* model. People understand the difference between self-interest (maximization of utility) and values in an intuitive way. In the above-mentioned example, people improve their situation when they leave money for themselves. On the other hand, behaving according to the values (giving money back) seems to make their situation worse. This way of thinking is due to the fact that people do not want to be 'good' because this gives them some profit (utility). Moreover, they want to be 'good' because being 'good' has some intrinsic value independent of utility. In order to explain the difference between self-interest and values, it is useful to use the concepts of 'sympathy'⁵ and 'commitment' which were proposed by Sen (1977, p. 326):

The former corresponds to the case in which the concern for others directly affects one's own welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment.

Many economists have a problem to notice the above-discussed difference. It is visible when altruistic and self-interest motivations are put into the black box of utility. It means that economists do not have tools to notice the difference between these two distinctive motivations (both can maximize utility in value-free economics). It causes that economists look at human behaviour in a particular way. For almost 150 years (marginal revolution in 1870), economists have assumed that human beings are predominantly driven by self-interest. This assumption together with methodological individualism is one of the cornerstones of neoclassical economics. However, it has not always been like that. Adam Smith perceived people in a complex way. The majority of economists refers only to *An inquiry into the nature and causes of the wealth of nations* (1776) and they forget about *Theory of moral sentiments* (1759) which had been written 16 years before the 'economic bible.' In this book, Smith wrote about the "altruistic side" of human beings. He wrote about an impartial spectator, social norms and sympathy.⁶ Smith considered himself as a moral philosopher (McCloskey, 2008; Wells, 2013), but because of different reasons his moral philosophy has been quite forgotten. One of these reasons was that Smith's legacy was used by many economists for their own purposes (Sen, 2010). This leads to a situation in which nowadays the Scottish thinker is associated predominately with both concepts: self-interest and invisible hand (Ostapiuk, 2017b).

⁵ Sympathy is one of the most important concepts in *Theory of moral sentiments* (Smith, 1759). It is worth noticing that Sen is a great admirer of Smith and he often uses insights of Scottish thinker

⁶ The differences between the two books befuddled many economists. The contradiction was called *das adam smith problem* (this contradiction is false, see more (Dixon, & Wilson, 2006; Ostapiuk, 2017b)).

It is essential to understand how neoclassical economists treat Smith, because it tells us a lot about neoclassical economics. It assumes that people are self-interested, economics does not deal with people's ends (value-free economics), and people always maximize their utility (which can mean everything). Etzioni (1988) wrote an extensive book where he criticized the simplification of these assumptions. The purpose of this subchapter is less extensive. A few examples will be presented, and they should convince the reader that it is impossible not to distinguish between human motivations.

To begin with, it is essential to understand that value-free economics perceived self-interest very broadly. This understanding of self-interest is the result of the criticism of homo economics concept. Frequently the other social sciences treat *homo economicus* as a straw man (Dixon, & Wilson, 2012; Urbina & Ruiz-Villaverde, 2019). The most controversial assumption of economics is that people behave in accordance with their self-interest.⁷ Many opponents of *homo economicus* ask why people give their organs to those who they do not even know, why people give tips in a roadside restaurant or why people give presents. Value-free economics had to expand the notion of self-interest to its extremes in order to answer the above-mentioned questions. One anecdote about Hobbes can help us to understand how broadly self-interest is perceived by value-free economics.

Apparently, Hobbes was walking in London with a clergyman, and suddenly they spotted a beggar. Hobbes gave alms to the beggar and this astonished the clergyman. He said: 'You helped this poor man, so, your theory about an egoistic man has fallen down.' Hobbes responded: 'No,' and he added that 'It proves my point that I am an egoist because looking at this poor man makes me feel unhappy and when I gave him money I felt better' (MacIntyre, 1990, p. 135). In philosophy this approach is called psychological egoism and it assumes that people always behave in accordance with their self-interest. Even if they help and call it altruism it is only a disguise for their self-interest. Value-free economics has used this reasoning for many years. It is possible to find an analogy in this context with utility maximization. In psychological egoism, every behaviour can be explained by an egoistic motivation. One just needs to look for it deep enough. The economic approach assumes that people always maximize their utility (perceived *ad libitum*) and *ex-post* it is possible to consider every behaviour as the result of utility maximization. Which implications does it have? For example, value-free economics cannot differentiate between the soldier who jumps on a grenade to save his comrades from the soldier who

⁷ The criticism started with Mill and his conception of *homo economicus*, and it lasts to this day (Boulding, 1969; Chmielewski, 2011; Cohen, 2014; Etzioni, 1988.; Frey, & Oberholzer-Gee, 1997; Hausman, & McPherson, 2006; Hodgson, 2012a; Kahneman et al., 1986; Ostapiuk, 2017b; Reiss, 2013; Sen, 1987; Thaler, 2015).

pushes another soldier on a grenade to save himself. This relativistic approach (value-free) together with the assumption of a human being as a self-interest entity causes that economics looks at reality in a very particular way.

Now, we can come back to the questions connected with donating organs to people that we do not know, giving tips in restaurants or giving presents. The answer is simple for value-free economics. People do it, because it gives them utility. What is more, people choose something, because it gives them the most utility. According to value-free economics, if one gives a kidney to the stranger, it is only because it gives utility to a person. Becker understood that people have different motivations (e.g., love, envy, brotherhood). However, his approach does not take this difference into consideration. People just maximize their utility. Thus, it means that if one gives a kidney to some stranger, one will do it, because it will give one utility (pleasure from being a good person).

This broad perception of self-interest raised many questions. It was Jean-Jacques Rousseau who presented some arguments against psychological egoism, and his argumentation can be used against the wide perception of self-interest in economic approach as well. If people can take their own interest and the interest of the others into consideration (even if in the end they will behave in accordance with their self-interest), it means that people must be somehow connected with the other people. At least, to the extent when the interest of others is depicted as the alternative to their own self-interest (MacIntyre, 1990, p. 186). The problem with value-free economics is that it is solely focused on the consequences of people's actions. In the end, the feelings of love, friendship and empathy give people pleasure (utility). However, the question arises whether they are just the motivations for our actions or the results of them. Do people want happiness for their children, because it gives happiness, or if this happiness is the result of their love for children? Do people give presents because giving presents makes them happy or is it the result of the bonds with the other people? We do not need to answer these questions, because as it was noticed by Rousseau, it is enough to incorporate others into our happiness to make the conception of egoism too broad (see also (Ostapiuk, 2017b)).

This lack of interest in human motivations causes some problems which were not considered by Becker. He gives the example of a family where the wife's behaviour is transformed into the altruistic one due to self-interest (Becker, 1974). In this example, the husband likes to read books in his bed before he falls asleep. His wife does not like this, because it disturbs her sleep. However, the wife does not forbid him to read, because she knows that a happy husband (in the higher level of utility) will compensate her 'sacrifice' with interest. Her behaviour is the same as the wife who really cares about her husband but their motivations are different. What is the most important here is this husband prefers to be loved rather than not to be loved. He wishes that his wife permits him to read,

because it gives utility him not her. It means that choice alone (consequence) is not enough and that motivations do matter for people.

Amartya Sen is the most recognizable authority who argues that economics cannot be only focused on the act of choice. The background is important as well. Sen gives the example of a hungry man to underline this fact (Sen, 1999, p. 75). In the first case he is hungry, because he does not have food, whereas in the second case he is hungry because of religious reasons (fast). Value-free economics cannot see the difference between these two behaviours, because it is not interested in human motivations. It is only interested in choice alone.

The next problem of economic approach is its understanding of rationality. We can realize how the protective belt of economic approach works on the example of *A theory of rational addiction*. There, Becker and Murphy (1988, p. 675) claim that:

Yet, as the title of our paper indicates, we claim that addictions, even strong ones, are usually rational in the sense of involving forward-looking maximization with stable preferences. Our claim is even stronger: a rational framework permits new insights into addictive behavior.

Drugs are the topic where Becker pushes rational choice theory to its logical extreme. He did it, because he wanted to know if his approach would work. The point of the article is that addicted people are not irrational and their behaviours are not anomalies at all. What is more, Becker and Murphy argue that taking drugs maximize addict's utility. This is possible, because Becker and Murphy use the instrumental rationality: "this paper relies on a weak concept of rationality that does not rule out strong discounts of future events. The consumers in our model become more and more myopic as time preference for the present (a) gets larger" (Becker, & Murphy, 1988, p. 683). This means that taking drugs is rational, because pleasure from drugs now is so immense that it outweighs future costs. Therefore, taking drugs can maximize people's utility. It looks strange that being addicted to heroin is rational if one looks at the results of addiction. However, it is not a problem, because, for Becker, economics does not assume what is good or bad for people. We just assume that individuals maximize their utility in the moment of choice. In the case of drugs, individuals make decisions between present and future utility. Becker assumes that weakness of will (*akrasia*) does not occur and people somehow compare their long-term goals (e.g., health) with the short-term pleasures (e.g., smoking). We (as economists) do not know how they do it (we are not interested in motivations). We simply assume that they can do this (as if assumption).

The belief that people maximize their utility is very strong in the economic approach. Even if people regret some decisions made on the spur of the moment, it does not mean that they do not choose what they want. Becker and Murphy (1988, p. 693) write that:

The claims of some heavy drinkers and smokers that they want to but cannot end their addictions seem to us no different from the claims of single persons that they want to but are unable to marry or from the claims of disorganized persons that they want to become better organized. What these claims mean is that a person will make certain changes—for example, marry or stop smoking – when he finds a way to raise long-term benefits sufficiently above the short-term costs of adjustment.

For Becker people's declarations do not mean anything. If abstinence from smoking gives one more utility than smoking, one will quit cigarettes. Ultimately, in the economic approach it is one's deeds, not one's words that matter most.

The problem with this understanding of rationality is that there exist many examples when it is strange to assume that people do something, because they really want it. On some level when we look at how much pleasure can be obtained from using drugs, then we can agree with Becker that seemingly irrational behaviour as taking drugs can maximize utility. However, the question is if this understanding of rationality makes any sense. In the economic approach, ex-alcoholic who does not want to drink, because she knows where it can take her, is rational when she comes back to drinking. She does not want to drink and after going on a binge, she feels awful about her behaviour. Is it possible to decide if the temporary need to relapse to alcohol is more important than the goals and human motivations? Becker assumes that human beings can integrate these conflicting motivations and bring them down to a common denominator of utility. This understanding of rationality gives rise to some absurd conclusions:

Imagine the situation where, owing to the shipwreck a man lies on a small boat in the middle of the ocean, and he thinks about drinking water from the sea. He knows that he should not drink this water and it can only make him feel bad. Despite this knowledge, he succumbs to the temptation and drinks it. He gets sick and dies afterwards. Is it reasonable to say that this person behaves rationally? Is impulse, second utility enough to call some behaviour rational? (Ostapiuk, 2019b, p. 643).

The above-mentioned examples raise the question if we should not make a distinction between reason and biological needs. Instrumental rationality does not differentiate between individuals' behaviours and it can mean anything which leads to a tautology. Obviously, the prior examples are extreme; but the questions that result from them concern the majority of human behaviours. Individuals frequently choose between now and later. Consume now or save for later: have fun now or work hard now and have much more fun later. Becker and neoclassical economists assume that people always choose what is good for them. However, after examination of hyperbolic discounting and findings of behavioural economics it will be difficult to agree with this assumption.

Individuals' behaviour is not always time consistent, which is assumed by neoclassical economics. Therefore, it is very problematic to compare present

choices with future ones. It is true that neoclassical economists are focused on human choices in a particular moment. However, they knew that choices now and choices in the future differ from each other quite significantly. More than 100 years ago Jevons, one of the fathers of Marginal Revolution, wrote:

To secure a maximum of benefit in life, all future events, all future pleasures or pains, should act upon us with the same force as if they were present, allowance being made for their uncertainty. The factor expressing the effect of remoteness should, in short, always be unity, so that time should have no influence. But no human mind is constituted in this perfect way: a future feeling is always less influential than a present one. (Jevons, 1911, pp. 72, 73)

Economists have tried to resolve the problem of different choices over time. They have, for instance, acknowledged that present choices are valued more highly than later choices and this is why they discount future utility. Samuelson (1937) introduced the discounted utility model. The reference to the financial market terminology (discount) is not accidental. In neoclassical economics individuals are presumed to assess and enumerate utility that they would receive in the future. In his short paper (seven pages long) Samuelson spends the two last pages to point out some possible problems with his model that he calls "serious limitations" (Samuelson, 1937, p. 159). One of his concerns is whether people discount future at rates that vary over time, for instance, they may change their minds as time moves forward. Therefore, it bothered Samuelson that people might display time-inconsistent preferences. He was aware that such behaviours existed. For example, he wrote about purchasing total life insurance as a compulsory savings measure. However, he did not investigate the problems regarding inconsistency further. Economists followed his suit and the discounted utility model became a standard model of intertemporal choice (Thaler, 2015, p. 101). Modigliani (1966) is the best example of implementing the discounted utility model. He based his model on an individual's total income and his theory is called the 'life-cycle hypothesis.' In this theory, people are rational, and they are presumed to be able to make a plan when they are young about how to smooth their consumption over their lifetime. What is more, this life-cycle hypothesis does not only assume that people are able to make all calculations (with rational expectations) about how long they will live, how much they will make, etc., they also possess self-control which is required to achieve the optimal plan. It is not surprising that with time, economics dealt with the increasing number of cases which show that individuals cannot discount future utility according to an exponential rate which is a result of present bias. In fact, people try to discount future utility but frequently they fail to predict what will improve their well-being in the future (Gilbert, 2007). Moreover, the complications with discounting utility are not only the ordinary anomalies but they indicate a systematic problem which suggests that the foundations of economic man – stable preference – cannot be fulfilled.

One of the problems with the evaluation of the future is called present bias, which is one of the main reasons for hyperbolic discounting (Ainslie, 1975, 2001; Loewenstein, & Elster, 1992; O'Donoghue, & Rabin, 1999). To illustrate what present bias is, consider the following example. People can choose \$ 100 in 30 days or \$ 105 in 31 days. The majority of them choose the second option. However, if they must choose between \$ 100 today or \$ 105 tomorrow, almost everyone will pick the first option. The difference between the choices is not a problem for economics, because the discounted utility model assumes that the future is less important than the present (Samuelson, 1937). Therefore, waiting one day after 30 days to get \$ 5 is easier than waiting one day from today. However, as it turns out, people do not discount the future according to an exponential rate. They opt for \$ 105 in 31 days but when the 30th passes and they have the choice between \$ 100 now or \$ 105 tomorrow, they will switch their preferences to \$ 100 now. This behaviour is inconsistent with the long-term preferences and it means that individuals over-value present choices. All along, they want some choice more than the other but when selection arrives then their preferences will change. For Becker and neoclassical economics present bias is not a problem, because in the end individuals make choices and their cognitive biases are the reason why rationality is not perfect. The issue here is that Becker and neoclassical economists assume that people always make the best choices. In the context of present bias it means that economics opt for an agent who chooses simple pleasure over long term goals.

People, however, know about the problems with decisions over time. Odysseus is depicted as an illustrative example of human's conflict between the long-term goals and the short-term drives (see (Elster, 1987)). He was supposed to float over the rocky islands which were inhabited by sirens. Their singing ensnares many sailors and as a result, they hit the rocks. Ulysses knew about sirens' appeal and he decided to protect himself. In order to do it, he commanded some deckhands to tie him to the mast and to put wax in their ears. Thanks to it, they would not hear the sirens' singing. Due to these precautionary measures, the ship has floated luckily and after that Odysseus boasted that he heard the beautiful sirens' singing.

This behaviour of mythical hero is known as 'precommitment strategy.' People use this strategy of their own volition to remove options from which they choose, because they are afraid that they succumb to the temptations when the choice comes. For Schelling this behaviour was really disturbing when he referred it to the neoclassical economics and revealed preference theory. This is why he decided to inspect it (Schelling, 1978, 1980, 1984, 1985, 1996). For him choices between long-term goals and short-term pleasures are not anomalies but they are common to human reasoning and judgmental processes. Precommitment strategy is one of the reasons why Schelling believes that people are not homogenous agents. Therefore, it is impossible to indicate one self with well-

-defined preferences but we need to indicate multiple selves with a different set of preferences which clash with each other. Due to these differences, it is impossible to compare utility over time:

Each self is a set of values; and though the selves share most of those values, on the particular issues on which they differ fundamentally there doesn't seem to be any way to compare their utility increments and to determine which behavior maximizes their collective utility. (Schelling, 1984, p. 8)

We cannot use utility measurements, because human beings have two different selves and they have diverse preferences.⁸ Thus, economics should handle these selves separately. In Chapter 4 this topic will be examined more thoroughly.

2.5. Becker's unsuccessful escape from tautology

The goal of the analysis of the economic approach was to indicate that thanks to the axiomatic assumptions, value-free economics could absorb every criticism from the other social sciences and paradigm remain unchanged. However, the extension of the protective belt leads to many problems. Because of the extension of perception of rationality and utility, Becker could defend hard core assumptions but at a price of becoming a tautology.

Becker did not want his theory to be a tautology and this why he introduces stable preferences. In this book this assumption is not considered as the hard core but as the negative heuristic (the same understanding in Blaug 1992, p. 221). Becker had two main reasons for stable preferences. Firstly, he is the successor of the positivistic approach and value-free economics. Therefore, he believes that we cannot scientifically discuss tastes and preferences. The unwillingness to investigate human motives is visible when we look at the title of one of the most significant Becker's articles *De gustibus non est disputandum* (there is no accounting for tastes). In this article, Becker underlines the meaning of stable preferences: "One does not argue over tastes for the same reason one does not argue over the Rocky Mountains - both are there, will be there the next year, too, and are the same to all men" (Stigler, & Becker, 1977, p. 76). Becker was aware that he simplified reality by assuming that stable preferences exist. However, this simplification was necessary, because he did not want to be entangled in the infeasible analysis of human motivation. For him the preferences cannot be empirically tested (see (Caplan, 2003)). Moreover, they provide "endless degrees of freedom" (Stigler, & Becker, 1977, p. 89). Becker makes even stronger statements about preferences. He writes that "no significant behavior has been illuminated by assumptions of differences in tastes" (Stigler, & Becker, 1977,

⁸ More about multiple selves in: (Ainslie, 2001; Cowen, 1991; Davis, 2003, 2010; Elster, 1987; Frederick, Loewenstein, & O'Donoghue, 2002; Fudenberg, & Levine, 2006; Heilman, 2010; Read, 2006; Strotz, 1955; Thaler, & Shefrin, 1981).

p. 89). Becker did not only state that economics should not analyze preferences. He concluded that the economic approach does not need any information about tastes to work.

The second reason behind stable preferences is the need for predictive power. Becker (1976, p. 5) writes that:

The assumption of stable preferences provides a stable foundation for generating predictions about responses to various changes, and prevents the analyst from succumbing to the temptation of simply postulating the required shift in preferences to 'explain' all apparent contradictions to his predictions.

To show why the assumption of stable preferences is so essential, Becker examines the opposite assumption – that tastes are variable and subject to change. Then, in the case of some anomalous behaviour the shift in taste would resolve the issues, leaving no other eventuality to be explored and no more to be done. Therefore, a shift in taste could explain any anomalous behaviour. This means that the assumption of changing tastes explains too much and “a theory that explains everything, explains nothing” (quote attributed to Popper). Becker did not want his economic approach to be pseudo-sciences like, in his opinion, theories of Adler, Freud, and Marx. Therefore, he had to assume stable preferences. For him this assumption resolves the previous problems of neoclassical economics:

On the traditional view, an explanation of economic phenomena that reaches a difference in tastes between people or times is the terminus of the argument [...] On our preferred interpretation, one never reaches this impasse: the economist continues to search for differences in prices or incomes to explain any difference or changes in behaviour. (Stigler, & Becker, 1977, p. 76)

However, the assumption of stable preferences has one simple flaw. It contradicts observable facts. Preferences change over time and it makes the economic approach unable to deal with choices over time. Each time when people choose between now and future (which is a common dilemma of human behaviour), the economic approach will not be able to indicate what people should choose (what is the best for them). Of course, it seems not to be a big issue for the value-free approach but this objectivity is only illusory. In the economic approach, it is normatively assumed that people do what gives them the most utility. Therefore, what people do is the best for them (revealed preference). The phenomena of present bias and hyperbolic discounting are so ubiquitous that they divide human being into multiple selves. These selves have absolutely different preferences. Having this in mind, it seems justified to make a distinction between 'short-term human' and 'long-term human' (Ostapiuk, 2019b).⁹ Value-free economics does not decide what people should

⁹ The former is driven by emotions and temptations which are predominately connected with short-term pleasures. The latter has long-term values and goals. “Long-term human” can be identified with reason and self-reflection.

value and what gives them the most utility. However, this value-freeness is an illusion. When economists assume that people always maximize their utility in reality it means that economists unconsciously support the 'short-term human.' It happens due to hyperbolic discounting which demonstrates that people are too much (irrationally) focused on the present and choose short-term pleasures.

The economic approach cannot admit that individuals' preferences change over time, because without the assumption of stable preference, the economic approach will not have any predictive power. Value-freeness of the economic approach means that economists cannot say anything about human preferences before the choice is made. Before the choice it can only be stated that people will maximize their utility and they will behave in a rational way. However, due to the tautologicality of these statements it does not mean anything (or it does mean everything). It is just a more sophisticated way of saying: people do something, because they do something. Without the assumptions of stable preferences, the economic approach could only analyse choices *post factum*. For example, Mark decides to buy this car, because it maximizes his utility, but we could not predict what he will do.

In the end, the escape from tautology did not work out and this problem is indicated by Blaug, who suggested that it is hard to find novel facts in Becker's writings, or even that they do not exist. He pointed out that: "Becker's writings lend themselves all too easily to caricature because they employ a cumbersome apparatus to produce implications that are sometimes obvious, if not banal" (Blaug, 1992, p. 223). For Blaug, Becker's approach is mainly based on dressing up intuitive or well-known conjectures about human behaviour with fancy and unwieldy formal mathematics which are called predictions. Although these predictions are described as testable and falsifiable, the economic approach is often guilty of being so general that it is "compatible with almost any finding" (Blaug, 1992, p. 224). Blaug noticed that Becker was always able to add some additional *ad hoc* assumption (like instrumental rationality and utility *ad libitum*) that would explain every anomaly. Blaug criticized this approach and called it "ad hocery" (Blaug, 1992, p. 250). The previously mentioned problems demonstrate that we can perceive the economic approach as degenerative programme. The main reason for this is the fact that new auxiliary hypotheses are not proposed to deal with new facts better but only to defend hard core (similar conclusions we can find in (Blaug, 1992), but our analyses and reasoning differ significantly).¹⁰

¹⁰ The analysis of economic approach is only a part of Blaug's book (1992, pp. 206–228). He analyses 'new home economics' (Becker, 1964), whereas the focus in this book is put on theory of rational addiction (Becker, & Murphy, 1988). We also identify different hard cores of economic approach (in this book rational behaviour is divided on maximisation of utility and instrumental rationality, whereas Blaug does not distinguish them). The analysis of economic approach conducted here is broader and it concerns economic imperialism is general, whereas Blaug is mainly focused on Becker's theory of family.

Of course, the economic approach is not as bad as it looks when one is only looking at its flaws. Ultimately, Becker is perceived as a maverick to whom such terms as 'originality' and 'creativity' fit like a glove. He initiated a few new areas in economics and sociology. However, it does not change the fact that the economic approach looks more like degenerative than progressive scientific programme. The reason for this situation is the fact that the economic approach needs to add endless *ad hoc* adjustments which take care of every anomaly. There were so many anomalies which treat 'hard cores' that now it is assumed that every choice maximizes utility and is rational. The reason for the expansion of the protective belt to tautologies was not to explain new facts better, but only to defend the hard cores. Of course, the economic approach makes a prediction about the future, but the question is whether it is possible to make good predictions when the assumption on which the model is based (stable preferences) does not correspond with observable reality. Becker and Friedman argued that it is possible, but it is highly questionable if these kinds of predictions have some value (non-truistic predictions).

2.6. Conclusions

Having in mind the section on Becker's methodology, it seems justified to state that his economic approach is based in many ways on instrumentalism where Truth is not considered as a necessary ingredient. A scientific theory must be useful, and the ability to predict new fact is one of the criteria of usefulness. The essence of instrumentalism is presented by Boland who writes:

When we take our television set to the repair man, we do not usually think it is necessary to quiz the repair man about his understanding of electromagnetics or quantum physics. For our purposes, it can be quite adequate for him to believe that, for example, there are little green men in those tubes or transistors and that the only problem is that one of the little green men is dead. As long as the tube or transistor with the dead little green man is replaced and our television set subsequently works, all is well. (Boland, 1997, p. 51)

Maybe we do not need to know the truth in the above-mentioned case but science is not like repairing of a TV set. For Becker what really matters is predictive power and explanatory power is not essential. Moreover, the economic approach works better without true (corresponding with reality) assumptions. Philosopher Ernest Gellner shows what happens when one is not interested in the real world:

When I was a boy, I had a clock with a pendulum which could be lifted off. I found that the clock went very much faster without the pendulum. If the main purpose of a clock is to go, the clock was the better for losing its pendulum. True, it could no longer tell the time, but that did not matter if one could teach oneself to be indifferent to the passage of time. (Gellner, 1960, p. 15)

Gellner's part concerns linguistic philosophy but the moral of the story serve economics as well. In our story the economic approach, which cares only about prediction, and not about the world, "is like the boy who preferred the clock without the pendulum, because, although it no longer told the time, it went more easily than before and at a more exhilarating pace" (Gellner, 1960 p. 15).

It is true that Becker's economic approach logically works, deals with every anomaly, and economists can use this approach in different fields with success (as in the case of economic imperialism; see (Lazear, 2000; Mäki; 2009a; Mäki et al., 2017)). However, the question is if the economic approach can be improved, or is there no need for a better theory? The economic approach looks perfect and it seems to persuade many economists that there is no need for some different approach. However, when one stops believing in one of its assumptions, then the whole theory falls down. This is why the economic approach is reluctant to methodological pluralism or any competitive theory. The winner takes all. Therefore, either one accepts every part of the economic approach, or one needs to look for something else. Many economists do not want to look for something different, because when they dismiss value-free economics, then Pandora's box will open and they will need to study normative approaches. It would mean the necessity of analysing human goals and motivations as well as deciding normatively what is good for people. As one can imagine, it would diminish the actual effectiveness of economics which would be more similar to the other social sciences. Many economists also do not want change, because they believe in an unbridgeable normative/positive dichotomy. Therefore, they consider normative approaches as unscientific. On the other hand, they want to be as objective as it is possible and objectivity means for them an exclusively positive approach. Moreover, positivism has a huge influence on economics and positive approach started to be identified with scientific one (economists as engineers).

In the previous chapter, it was argued that perceiving positive approach as scientific and normative one as the unscientific method is unjustified. However, still many economists believe in this unbridgeable dichotomy. This is the reason why economics of happiness which studies what gives people well-being is treated by many as a fad. However, as it was indicated before, neoclassical economics is not as objective as it is commonly thought. Pareto optimum and the assumption that maximization of utility is something desirable are normative statements. The problem here is that many economists consider these statements true without realizing that they only believe that they are true (Sedláček, 2011). Still, they assume that people maximize their utility and that effectiveness is something good *per se*. Moreover, they often do not realize that positive economic statements are theory-laden and they are not so different from normative statements. On the other hand, normative statements are stigmatized as unscientific and they are identified with ethical judgments. Therefore,

in accordance with a purely positive approach, we cannot discuss them. Of course, one will not find absolute truth and establish what improves human well-being in every situation but it should not be a problem. We know that project of finding the Truth in science did not work out (Feyerabend, 1993; Kuhn, 1962; Lakatos, 1980; Rorty, 1979) but it does not mean that there is nothing to be done. For example, we can discuss and reason why some theories are better than the others. This process will be investigated in Chapter 5. This discussion is vital and indispensable for the development of economics as a science. There are no universal criteria for progress in social sciences. Following Kuhn (1962), a scientific community decides which theory is better. The economic approach is the dead end in this context, because it does not leave the room for discussion (you need to accept all premises) and without discussion the improvement is impossible.

The goal of this chapter was to demonstrate that value-freeness does not work well for economics in terms of descriptive and predictive power when it comes to the intertemporal choices. The economic approach was used as an example of what happens when one pushes hard core assumptions as instrumental rationality and maximization of utility to the extreme. The economic approach can be perceived as the embodiment of value-freeness and we can observe what problems it creates. From the perspective of this book, choices over time are the most important ones. Neoclassical economics assumes that weakness of will does not occur, because people always choose the best option. This assumption is treated as positive one and it does not represent reality correctly. Due to present bias, we know that people follow short-term preferences; and they underestimate long-term ones even if they want to stick to the longer perspective (e.g., (Strotz, 1955)). Therefore, when economics assumes that what people do is the best for them, then it unconsciously supports 'short-term human' (Ostapiuk, 2019b). If we agree that people have multiple selves with different preferences that are irreconcilable by utility measures, we will need to know what improves people well-being in order to choose one of the selves. It eventually implies that investigations on well-being can be the ultimate goal for economics. It would be very hard to reject the idea that the ultimate goal of life is well-being. The question that seems to be the most important is the question how to achieve well-being to which economists answer: you decide people.

One of the goals of this chapter was to demonstrate on the example of Becker's economic approach that axiomaticity of value-free economics result in the situation in which paradigm in economics has not changed despite being the subject of relentless criticism. It is the result of the extension of the protective belt in which the concepts of rationality and utility can mean everything. It results in tautologization of these terms. Despite its effectiveness and deductive consistency, value-free economics faces crucial problems (e.g., choices over time, well-being).

It goes without saying that Becker's economic approach and value-free economics have their advantages and they were used effectively by economists. It has not been the goal of the author to come up with a conclusion that there is the need to throw it away. The conclusions from this section are more modest. Due to the problems with value-free economics analysed in this and the previous chapter, we should consider economics of happiness, mostly focused on studying well-being, as a full-fledged scientific programme that deserves the attention of the economist's community. The analysis of economics of happiness is one of the main goals of the next chapter.

Chapter 3

Economics of happiness

*Every pitifulest whipster that walks within a skin
has had his head filled with the notion that he is, shall be,
or by all human and divine laws ought to be, 'happy'*

(Carlyle, 1843, p. 192)

3.1. Introduction

In this chapter, the author will check whether economics of happiness is a progressive programme which can resolve the problems that value-free economics faces (well-being and choices over time). Moreover, the author will analyse objective well-being approaches which aim to establish goals which are independent of happiness and utility which are used in economics of happiness. However, objective well-being approaches will be analysed in the second part of this chapter. Now, our attention is focused on the economics of happiness. The author will not provide a summary of economics of happiness (it was done by (Bruni, & Porta, 2007; Frey, 2010, 2018; Frey, & Stutzer, 2002a, 2002b)). The goal of this chapter is to outline the differences between value-free economics (revealed preference theory) and economics of happiness (subjective well-being). The latter deals with human goals where happiness is the ultimate one. In order to know what gives individuals the most happiness, researchers measure happiness by SWB measures (subjective well-being). Although it can be a step in the right direction for economics in terms of explanatory and predictive power, we need to check whether subjective well-being can replace revealed preference theory, or we need to search for some objective approaches. Before answering these questions, some basic knowledge about the economics of happiness must be presented.

Economics of happiness came to life in 1974 when Easterlin presented the paradox which later started to be called 'Easterlin Paradox' (Blanchflower, & Oswald, 2004; Diener, & Oishi, 2000; Easterlin, 1974, 1995, 2001; Kenny, 1999) or 'Happiness Paradox' (Pugno, 2004a, 2007). Easterlin noticed that in several countries real income had risen significantly since the Second World War, but the self-reported subjective well-being of these populations has not increased.

For example, in the United States, between 1946 and 1991, *per capita* income increased by a factor of 2.5 (from \$ 11,000 to \$ 27,000 in 1996 dollars). However, over the same period happiness lasted on the same level (Frey, 2010, pp. 11, 12). Easterlin pointed out that generally rich people are happier than poor people in the same country. However, on average people in more affluent countries are not necessarily happier than people in poorer ones. Moreover, rises in income over time do not increase happiness beyond a certain level. Easterlin's seminal paper (1974) started a big discussion on happiness and income that lasts to this day. In general, there is a consensus that income has a positive correlation with happiness, but after some level of income the correlation is weak (Hagerty, & Veenhoven, 2003; Stevenson, & Wolfers, 2008).¹ There are a few theories that explain why income after some level does not have a significant impact on human happiness (1 - hedonic treadmill theory, 2 - law of diminishing returns, 3 - happiness set point theory), but they will not be scrutinized here. Now it is important only to underline the relativism of happiness and the fact that people are not able to make absolute judgments about happiness. Individuals constantly make comparisons with an environment ('keep up with the Joneses'), with the past, or with the expectations of the future. This is why Frey (2010, p. 30) indicates aspiration levels as the cause of Easterlin paradox.² This relativity is essential, because later this phenomenon will be presented as one of the main reasons for dismissing subjective well-being measurements.

The topic of income and happiness is not the only one in the economics of happiness. However, this issue attracts the most attention. This topic is so interesting, because the statement 'more money more happiness' is widely accepted among mainstream economists (Frey, & Stutzer, 2002b, p. 428). A higher income expands people's opportunity set which means that more services and goods can be consumed. The bizarre (in economic theory) people who are not interested in more goods do not need to consume them. The unwanted surplus of goods can be managed without any additional costs. Therefore, economics takes for granted that happiness and income go together (Frey, & Stutzer, 2002b, p. 428). Some economists did not like this assumption and they openly criticized not only economic theory but a capitalistic system as well. Besides Easterlin, two other economists have had a huge influence on the economics of happiness. The first of them was Galbraith (1998). He criticized the capitalistic system in the book (among many others) entitled *Affluent society*. In his opinion, capitalism creates new needs which are limitless.

¹ Some scientists have been trying to establish the level of income after which people are not more happy, but there is no agreement about the exact number, because there is too many factors other than money which influence human happiness.

² For evidence from laboratory experiments about importance of aspiration levels, see (Mellers, 2000; Smith, Diener, & Wedell, 1989; Tversky, & Griffin, 1991).

Moreover, even if they are met, it does not necessarily give people satisfaction. The second economist Scitovsky (1976) showed that the increase in GDP does not make people happy. Therefore, economists need to be interested in philosophy and psychology, if they want to learn what brings happiness to people. More recently, Frank (1999) underlines that ever-increasing consumption and income do not guarantee people happiness and can even diminish it. This paradox is possible, because people measure their happiness in comparison with others (this phenomenon is known as ‘arms race’). So, if you get a new car, but your friends buy the same one, you will not be happier, because the social status is a zero-sum game.

Of course, economics of happiness cannot be simplified only to the search for a correlation between income and happiness. This approach measures what effects on happiness have things such as marriage, unemployment, freedom, democratic rights, watching TV, self-employment, inflation, and much more. Over the last years the research in happiness has attracted a lot of attention from the interdisciplinary community. Obviously, philosophers have been interested in happiness from the very beginning (e.g., Aristotle, Plato, Socrates, Epicure). However, the rest of the social scientists started to be interested in happiness very recently. Psychology started to deal with human happiness in the 1990s. Before that psychologists had been mainly focused on mental illness and the other problems which diminish human happiness. In the 1990s new approach occurred and was called positive psychology. The name was not accidental, because from that time psychologists have tried to find out what brings people happiness (Argyle, 1987; Csikszentmihalyi, 1990; Diener, 1984; Michalos, 1991; Myers, 1992; Nettle, 2005; Ryan, & Deci, 2001).

Economists earlier than psychologists started to investigate human happiness (Easterlin, 1974; Scitovsky, 1976). Although many economists noticed their work, they did not have many followers at the time. General interest among economists in the determinants and measurement of reported well-being started at the symposium in London in 1993. Later the proceedings from this meeting were published in *The Economic Journal* (e.g., (Frank, 1997; Ng, 1997; Oswald, 1997)) and other sources (e.g., (Clark, & Oswald, 1994, 1996)). In the late 1990s, economists started to publish numerous large-scale empirical analyses of the determinants of happiness in various countries and periods (Frey, 2010, p. 13). This interest in happiness is not an absolute novelty for economics. Before value-free economics and ordoliberal revolution happiness was one of the most critical issues for economists. Smith analyses the connection between wealth and happiness (see (Ashraf, Camerer, & Loewenstein, 2005)). For him the conviction that wealth can give happiness was false. However, he understands that in the end it is a good thing that people have this false conviction, because it is a fuel for the economy:

The pleasures of wealth and greatness, when considered in this complex view, strike the imagination as something grand and beautiful and noble, of which the attainment is well worth all the toil and anxiety which we are so apt to bestow upon it. And it is well that nature imposes upon us in this manner. It is this deception which rouses and keeps in continual motion the industry of mankind. It is this which first prompted them to cultivate the ground, to build houses, to found cities and commonwealths, and to invent and improve all the sciences and arts, which ennoble and embellish human life; which have entirely changed the whole face of the globe, have turned the rude forests of nature into agreeable and fertile plains, and made the trackless and barren ocean a new fund of subsistence, and the great high road of communication to the different nations of the earth. (Smith, 1759, IV.I.9–10)

He stated many times in *Theory of moral sentiments* that people find real happiness not when they meet self-interest motivations but when they behave morally and when they live in harmony with the others (Ostapiuk, 2017b). Smith (1759, VI.II.22) famously writes that:

No benevolent man ever lost altogether the fruits of his benevolence. If he does not always gather them from the persons from whom he ought to have gathered them, he seldom fails to gather them, and with a tenfold increase, from other people. Kindness is the parent of kindness; and if to be beloved by our brethren be the great object of our ambition, the surest way of obtaining it is, by our conduct to show that we really love them.

However, this guidance came rather from philosophy. The goal of economics at that time was to give people the possibility to accumulate more wealth. This approach was essential for the other classical economists. J.S. Mill established the subject of economics (at least for some time) when he wrote:

What is now commonly understood by the term ‘Political Economy’ is not the science of speculative politics, but a branch of that science. It does not treat of the whole of man’s nature as modified by the social state, nor of the whole conduct of man in society. It is concerned with him solely as a being who desires to possess wealth, and who is capable of judging of the comparative efficiency of means for obtaining that end. It predicts only such of the phenomena of the social state as take place in consequence of the pursuit of wealth. (Mill, 1874, p. 137)

Also Marshall (1920) perceived the goal of economics as the pursuit of wealth but he was more specific, because he wrote about material wealth.

This subject matter of economics has changed due to the processes connected with Pareto, Robbins, and Samuelson, which were presented in Chapter 1. It was indicated that value-free economics moved from assessing utility (cardinal approach) towards axiomatic assumptions in which people do what is the best for them (revealed preference theory). These seemingly ‘objective’ assumptions created a vacuum, because value-free economists did not deal with happiness, and now different researchers try to fulfil this vacuum. Moreover,

the assumptions that people make the best decisions and maximize their utility attracted much criticism. Economics of happiness tries to show that due to the bounded rationality people do not make best choices. These attempts to establish and measure utility are not something revolutionary in economics. Before Pareto revolution economists had tried to measure utility and they took human goals into consideration. Due to methodological weaknesses, it was not possible. Nowadays economics of happiness tries to present a coherent and scientific theory which will give economics the tools to deal with elusive happiness.

3.2. Subjective well-being

The main interest of economics of happiness is in measuring happiness. There exist many techniques for inquiring about individuals' well-being. They all give an indication of individuals' affective state (mood and emotions) or individual's evaluation of their life satisfaction and happiness. The umbrella term for all these measures is subjective well-being (Frey, 2010, p. 17). The most popular method which is used to measure human happiness relies on asking a representative sample of individuals about their general satisfaction with life. There exist many variants of this method. The example of a single-item question on a three-point scale came from the General Social Surveys (Davis, Smith, & Marsden, 2001): "Taken all together, how would you say things are these days - would you say that you are very happy, pretty happy, or not too happy?" In the World Values Survey (Inglehart, Basanez, Diez-Medrano, Halman, & Luijkx, 2000), life satisfaction is assessed on a scale from 1 (dissatisfied) to 10 (satisfied). People are asked: 'All things considered, how satisfied are you with your life as a whole these days?' The Euro-Barometer Surveys, covering all members of the European Union, ask an analogous question: "On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?" (Frey, 2010, p. 18). Among the multiple-item approaches, the most known one is the Satisfaction With Life Scale (Pavot, & Diener, 1993), which contains five questions rated on a scale from 1 to 7.³

The next method is called Experience Sampling Method (ESM). In this approach information on individual's experiences are collected in real time in people's natural environment (Csikszentmihalyi, & Hunter, 2003; Scollon, Prieto, & Diener, 2003). Moreover, a representative group of individuals gets a beeper or a hand-held computer. They are asked at random times to assess how happy they are and what are they doing. Thanks to this it is possible to assess what activities are pleasant/unpleasant for people, and at the end of the day we can calculate happiness by the aggregation of these instant statements or affects.

³ For a survey of various measures of subjective well-being, see (Andrews, & Robinson, 1991).

The next method is called Day Reconstruction Method (DRM). Questioners are asked to reconstruct the previous day by filling out a questionnaire. Firstly, they recall the activities undertaken the previous day by producing a sequence of episodes. Secondly, they describe each episode in detail. They identify what, when, where, and with whom the episode took place. Then, the respondents are asked to rate these episodes in terms of positive affect (warm/friendly, happy, enjoying myself) or negative affect (depressed/blue, frustrated/annoyed, hassled/pushed around, angry/hostile, criticized/put down, worried/anxious) (Frey, 2010, pp. 20, 21). In Table 1 the Day Reconstruction Method (DRM) is presented on the well-known example in the happiness literature of the women from Texas:

Table 1. The Day Reconstruction Method

| Happiness in different activities | Happiness (index) | Average hours per day |
|-----------------------------------|-------------------|-----------------------|
| Sex | 4.7 | 0.2 |
| Socialising after work | 4.1 | 1.1 |
| Dinner | 4.0 | 0.8 |
| Relaxing | 3.9 | 2.2 |
| Lunch | 3.9 | 0.6 |
| Exercising | 3.8 | 0.2 |
| Praying | 3.8 | 0.5 |
| Socialising at work | 3.8 | 1.1 |
| Watching TV | 3.6 | 2.2 |
| Phone at home | 3.5 | 0.9 |
| Napping | 3.3 | 0.9 |
| Cooking | 3.2 | 1.1 |
| Shopping | 3.2 | 0.4 |
| Computer at home | 3.1 | 0.5 |
| Housework | 3.0 | 1.1 |
| Childcare | 3.0 | 1.1 |
| Evening commute | 2.8 | 0.6 |
| Working | 2.7 | 6.9 |
| Morning commute | 2.0 | 0.4 |

Source: based on (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004, p. 1777).

The new method connected with the development of technology is brain imaging. This method relies on functional magnetic resonance imaging (fMRI) which tracks blood flow in the brain, using changes in magnetic properties due to blood oxygenation (Camerer, Loewenstein, & Prelec, 2005; Fehr, Fischbacher,

& Kosfeld, 2005; Zak, 2004). Happy people reveal a characteristic pattern of electrocortical activity. They manifest greater activity in the left prefrontal cortex than in the right (Davidson, 2003; Pugno, 2004b; Urry et al., 2004). However, at least for now this method is not very widespread among researchers. We need to wait for new experiments if we want to draw any significant conclusions from neuroeconomics in general (Braeutigam, 2005; Harrison, 2008; Marchionni, & Vromen, 2014; Rubinstein, 2008).

It goes without saying that subjective well-being (SWB) attracts much criticism. Especially surveys about general satisfaction with life have raised many questions but these questions relate to other methods as well. The skeptics ask if it is possible to believe in people's answers about subjective well-being if we know that emotions influence these answers. For example, someone who had a car accident before taking the survey can declare a lower level of happiness. Moreover, people are influenced by framing effect which means that those who see a documentary film about starvation in Zimbabwe would be more prone to declare a higher level of happiness, because they will compare their lives with poor lives of people inflicted by a starvation. The skeptics also ask if we can be sure that people tell the truth when they fill in a survey or they try to look happy if not in the eyes of others (anonymity of surveys), at least, in their own eyes. The questions also concern the possibility of assessing happiness on the simple numerical scale.

The researchers who analyse happiness have answers to all the above-mentioned questions. Although people are affected by positive or negative emotions caused by different experiences which had occurred before the survey, so many surveys were conducted that these emotions counterpoise each other. Moreover, there exists a strong correlation between the different surveys and their results. Most people like the same things, for example, spending time with friends and do not like things as commuting. Also, it was proven that different measures of subjective well-being correlate significantly with one another (Fordyce, 1988). The factor analyses of self-reports and non-self-reports of well-being revealed a single unitary framework underlying the measures suggesting their scientific validity (Sandvik, Diener, & Seidlitz, 1993). Despite the validity of subjective well-being measures, it is easy to disturb the results of some survey. In one experiment (Strack et al., 1988) people have to answer the analogous question. The first group was asked "How happy are you?" and later "How many dates did you have last month?" There was not any significant correlation between answers (0.12). The second group was asked first about dates, then if they are happy, and the correlation was significant (0.66). As it turned out, thinking about dates was an anchor which overwhelmingly influenced reports of subjective well-being. This experiment proves that it is easy to influence the results of an SWB survey. However, it does not diminish the validity of SWB surveys, because they are not designed to manipulate people's answers.

Probably the most substantial criticism of SWB measures concerns their simplicity. The assumptions that one simple question can provide a general assessment of emotional, social and mental health in comparison to hours of tests where the most respected clinical instruments in the field were used seems unrealistic (Fordyce, 1988). However, it works and the different surveys give the same results. Fordyce who was a pioneer in positive psychology considered this situation to be most fortuitous, 'since if great inconsistency in resultant findings occurred in the field, happiness studies would be in a thoroughly confused and confounded state' (Fordyce, 1986, p. 27). Apparently, "no matter how you decide to ask people how happy they are, the results are the same" (Fordyce, 1986, p. 27). When it comes to the validity of people's declaration about happiness, their answers correlate with the opinions of people closest to them. So, happy people are rated as happy by their family members and friends (Lepper, 1998; Sandvik et al., 1993) and by their spouses (Costa, & McCrae, 1988).

The last and most significant problem with SBW is its relativism. The first question that came to mind is how to measure different results. Is the difference between 4-5 the same as in 8-9, and if we can compare the level of happiness among individuals? So, if Mark declares that he feels eight on the 10 points scale, it means that he is happier than Tom who declared seven? Before admitting that this criticism is valid it must be stated that relativism of subjective well-being measures is inevitable. In the end, only people can decide how well they feel. However, due to many biases, people's declaration or even their belief that they feel happy does not necessarily mean that it is true. This point will be explained a few pages later when some arguments for objective goals are analysed.

It is true that subjective well-being data are based on individuals' judgments that are prone to a multitude of biases. Reported subjective well-being depends on the order of questions, the wording of questions, the scales applied or actual mood. However, the importance of these errors depends on how data is used. Frey (2010, p. 19) concludes:

Often, the main use of happiness measures is not to compare levels in an absolute sense but rather to seek to identify the determinants of happiness. For that purpose, it is not necessary to assume that reported subjective well-being is cardinally measurable or that it is interpersonally comparable.

This means that we should not expect from the economics of happiness to decide if Mark who declared eight in the survey is happier than Tom who declared seven. Economics of happiness has more humble goals such as an identification of the determinants of happiness. Of course, this means that economics of happiness has not given an ironclad argument on how to make interpersonal comparisons or assess cardinal utility. However, at least it gives a valuable scientific theory which draws us closer to the answer. Even though economics of happiness has not proven yet to be the only game in town when it came to

happiness, it seems better (in terms of explanation and accuracy) than value-free economics with its axiomatic assumptions about happiness. However, to state that with confidence, we need to analyse more thoroughly the revealed preference theory in the context of happiness.

3.3. The criticism of revealed preference theory

The theory of revealed preferences was analysed in Chapter 1. In this chapter, it is investigated in the context of happiness. Value-free economics assumes that choice tells us enough about preferences. This assumption is the result of an impossibility to test preferences empirically. Therefore, we need to wait and see what people choose to know their preferences. The problem with this idea is that neoclassical economics assumes that people satisfy their preferences (maximize utility) when they make a choice. Many economists do not treat this assumption as a normative one, see (Thaler, 1980, 2015). Many of them believe that in spite of the problems people are the best judges when it comes to preference satisfaction, e.g., (Friedman, 1962; Friedman, M., & Friedman R., 1980; Hayek, 1978) and most economists associated with Austrian school (Sugden, 2008b, 2010, 2018). Therefore, from this perspective, the only thing that economists can state is that people do something, because it satisfies their preferences. However, the implications of these assumptions have far-reaching consequences. Value-free economics assumes that what people do is the best for them. Therefore, what people do gives them happiness. These assumptions cause that value-free economics does not need to discuss happiness and human goals as it has a place in the other social sciences. This approach has consequences when it comes to knowing the reasons why people do something.

In a typical model of motivations a person has a desire Y , and if one believes that by doing act X , one can achieve Y , then one will choose X . Rational choice theory prescribes only the most effective ways to achieve given desires (Sugden, 1991). The only constraint is that desires must be consistent. Many philosophers are dissatisfied with such a structural definition of rationality. They want to know what desires are the best. So, in this context, the rational choice theory should answer questions such as: is it rational for a smoker to smoke, for an obese person to overeat, or for an employee to undersave (Read 2007)? Not answering these questions does not make value-free economics more scientific, because defining:

well-being or welfare exclusively in terms of subjective preference satisfaction, requires making a significant (and questionable) value-judgment, since well-being or welfare is also commonly defined in terms of a host of additional normative matters that the theory suppresses, such as justice, fairness, rights, liberty, and dignity. (Davis, 2005, p. 590)

Suppressing these values in the context of value-free economics means that all of them are put in the black box of utility where they lost their meaning. Moreover, a choice depends on opportunity set. Screpanti and Zamagni observe that it is epistemologically inconsistent to assume the compatibility between an individual's revealed preferences and the choices one can make:

preferences can be made operational [only] by means of a definition in terms of choice: the assertion 'the state x will be chosen by a subject if only x and y are available.' The doubt did not even cross the minds of Robbins and the other authors who followed this orientation that the definiens, as a conditional proposition, can perform its function only after the concept of preference has been defined. I may well prefer health to illness, but I certainly cannot choose to be well or ill. (Screpanti, & Zamagni 1993, p. 271)

The opportunity set has an enormous influence on choices that people made. Sen (2002, p. 171) writes that "We do not live in a world of a 'one-shot' choice". He indicates that without deeper knowledge about human motivations and goals sometimes people's preferences look inconsistent. The assumption of transitivity ($x > y$ and $y > z$ then $x > z$) is the foundation of rational choice theory. Sen gives an example to demonstrate that transitivity does not work in some situations:

Suppose the person faces a choice at a dinner table between having the last remaining apple in the fruit basket (y) and having nothing instead (x), forgoing the nice-looking apple. She decides to behave decently and picks nothing (x), rather than the one apple (y). If, instead, the basket had contained two apples, and she had encountered the choice between having nothing (x), having one nice apple (y) and having another nice one (z), she could reasonably enough choose one (y), without violating any rule of good behavior. The presence of another apple (z) make one of the two apples decently choosable, but this combination of choices would violate the standard consistency conditions, including Property a , even though there is nothing particularly inconsistent in the pair of choices. (Sen, 1993a, p. 501)

Sen's condition a is also known as Chernoff's Axiom and in general it is connected with axiom called 'the independence of irrelevant alternatives.' This axiom states that the unchosen alternative is unimportant. In the above-mentioned example with apples it means that a new alternative of another apple in the basket (z) should not change the choice from having nothing (x) to having one apple (y). Without consistency (transitivity) of preferences, the whole rational choice theory falls apart. To illustrate the importance of transitivity of preferences the anecdote which is attributed to philosopher Sidney Morgenbesser can be presented. After finishing dinner, Morgenbesser decides to order dessert. The waitress tells him he has two choices: apple pie and blueberry pie. Morgenbesser orders the apple pie. After a few minutes the waitress returns and says that they also have cherry pie at which point

Morgenbesser says “in that case I’ll have the blueberry pie” (Poundstone, 2008, p. 50).⁴ The unchosen option (cherry pie) causes that Morgenbesser changes his decision which does not make any sense, because in this case his preferences are inconsistent.

In Sen’s example a person is not inconsistent, because she follows the etiquette (does not want to look greedy) the whole time. Sen argues that we cannot know if some behaviour is inconsistent until we learn about one’s reason behind the choices. He writes that “we cannot determine whether the person is failing in any way without knowing what he is trying to do, that is, without knowing something external to the choice itself” (Sen, 2002, p. 130). Sen calls this something external “menu dependence” (Sen, 1997, p. 752). We would know that a person in the example with the apples is not inconsistent with her preferences if we knew about her motivations (not being greedy). Her preferences did not change, because of the *z* option. In the new environment (two apples) her preferences manifest themselves in a different choice.

Menu dependence causes that actual choice of the agent depends on the content of available alternatives. This makes internal consistency conditions meaningless, because choice alone does not take menu dependence into consideration. Sen writes that “I would argue that the philosophy of the revealed preference approach essentially underestimates the fact that man is a social animal and his choices are not rigidly bound to his own preferences only” (Sen, 1973, pp. 252, 253). The distinction between preference and choice is at the core of Sen’s critique of the revealed preference theory. This means that we need to know people’s motivation if we want to know about their preferences.

The next problem with revealed preferences is also connected with the opportunity set. In this case, however, the problem lies not in negligence of human motivations or preferences, but in the available options. Value-free economics assumes that people make the best choices for themselves. If they had not satisfied their preferences, they would not have chosen them in the first place. However, sometimes it is impossible to claim that some choices meet somebody’s preferences if we look at the available options. The example of Sophie’s choice proves this point. Sophie is the main character in William Styron’s *Sophie’s choice* (1992). She was a Polish woman who was relegated with her two small children (10-year boy and 7-year girl) to Auschwitz. For children being in concentration camp meant sure death. Therefore, Sophie tried to save their lives by seducing

⁴ Intransitivity of preferences is also known as ‘money-pump’. Hausman explains how it works: “Suppose for example I prefer *x* to *y* and *y* to *z* and *z* to *x*, and I start out possessing *z*. Then I should, in principle, be willing to pay fee for each of the following exchanges: trade *z* away for *y*, trade *y* away for *x*, and trade *x* away for *z*. I am then back when I started, except that I am poorer by the expense of the three fees. I have become a ‘money pump’” (Hausman, 1992, p. 16).

the commandant of camp for whom she worked as a translator. The plan did not work out and she was presented with the agonizing choice between saving only one child, whereas the second one would die being gassed. Of her two children, Sophie chose to sacrifice her seven-year-old daughter Eva, but the guilt haunted her for the rest of her life. Now the term 'Sophie's choice' is used in reference to a difficult situation in which a person must choose between two equally deserving alternatives (*Sophie's choice*, n.d.). In the context of revealed preferences it means that choice alone does tell everything about someone's preferences (of course, she did not want any of her children to die).

Another problem with revealed preferences is pointed out by the theory called 'second-order desires' (Frankfurt, 1971). This theory shows that the conception of revealed preferences is a simplification. Frankfurt (1971, p. 7) writes:

Besides wanting and choosing and being moved to do this or that, men may also want to have (or not to have) certain desires and motives. We are capable of wanting to be different, in their preferences and purposes, from what they are.

The easiest way to explain this concept is through the example of cigarettes. Having the choice to take or not to take cigarettes, the people who are addicted will choose cigarettes and this is their first-preference. However, if people had the choice, they would rather rank their preferences as no cigarettes > cigarettes than cigarettes > no cigarettes. This ability is called second-order desires. We should not perceive 'second-order desires' as the instability of preferences. It is rather individuals' preference ranking. We can also see the attempts to broaden the understanding of individuals' preferences by meta-rankings (Sen, 1977). The conception of second-order theory does not oppose Hume's thought that individuals cannot determine their preferences by reason ("reason is, and ought only to be the slave of the passions" (Hume, 1896, p. 283)). In the conception of second-order desires, people have defined preferences. However, due to reason they can change their preferences. The condition when people are tossed by their desires Frankfurt (1971) called "wanton". For Frankfurt wanton is the feature of young children and animals. In the second-order desires, consequences are not sufficient to show what preferences people really have. Therefore, it is crucial to analyse human's motivations and goals deeper.

The above-mentioned problems mainly concern philosophical issues. However, the revealed preference theory has even more problems with practical concerns. Some economists (e.g., (Friedman, 1953)) hold that economics is entirely descriptive and positive.⁵ Its only purpose is to predict human behaviour and nothing else. Neoclassical economists believe that the primary goal of the rational choice theory is to describe some regularities in human behaviour.

⁵ However, as was indicated before, Mäki (2009b) argues that position of Friedman is more complicated.

A theory works if it generates predictions that are consistent with observations. The problem of rational choice theory is that people are not always rational and they do not always behave in accordance with the economic theory. Moreover, irrational choices are not simple anomalies. People make constant mistakes due to cognitive biases (Thaler, 2015, 2016). Therefore, rational choice theory fails to predict human behaviour in some specific situations. Economics of happiness indicates that due to the bounded rationality, people do not choose the best options and they do not always choose things that will bring them happiness.⁶ The main mechanisms will be analysed below. They make revealed preference theory limited in cognitive terms.

Let us start with Kahneman's distinction between 'experienced utility' (cardinal utility) and 'decision utility', which is reflected in choices (revealed preferences) (Kahneman, 2011, p. 368). The experienced utility is connected with Bentham's idea to perceive utility as the feature of an object or an action in order to increase or decrease the overall happiness. Bentham's utility is logically separated from choices that are actually made. It means that nothing in the theory indicates that people will choose the best options for themselves. For example, smoking (decision utility) can give less total utility (experienced utility) than not smoking but be chosen anyway (Read, 2007). The definition of utility has changed over time. In value-free economics the separation between utility and choice disappeared. Decision utility does not say anything about hedonic experiences- what people do is the best for them and there is the end of a discussion. A choice needs only to be in accordance with the technical assumptions of instrumental rationality (consistency and completeness). The differences between experienced utility and decision utility do not mean that these concepts cannot coincide with each other. It is possible "if people want what they will enjoy, and enjoy what they chose for themselves" (Kahneman, 2011, p. 368). In general, revealed preference theory is based on the assumption that economic agents are rational. They are expected to know their preferences (tastes), both present and future. In other word, these preferences are well-informed. This enables them to make the best decisions which will maximize their utility (happiness). The problem is that individuals often do not choose the option which maximizes their experienced utility, and this makes them less happy.

Kahneman (2011, pp. 368, 369) formulates a puzzle in order to show the conflict between decision and experienced utility. In his puzzle individual receives one painful injection every day. There is no adaptation; the pain is the same every day. Kahneman asks, if people will attach the same value to reducing the number

⁶ Sometimes economics of happiness is mistaken for behavioural economics, because of the interest in bounded rationality. However, the goal of economics of happiness is different than the goal of behavioural economics. Happiness researchers focus on the fact that due to the bounded rationality people do not choose the options which bring them happiness.

of planned injections from 20 to 18 as from 6 to 4. He does not have doubts that people will pay more to reduce the number of injections by a third (from 6 to 4) than by one-tenth (from 20 to 18).⁷ The decision utility of avoiding two injections is higher in the first case than in the second. However, this difference is absurd, because two injections give the same amount of pain, and the number of previous injections does not matter at all. The example presented by Kahneman leads to two conclusions. Firstly, the experienced utility could be measured by the number of injections. Secondly, in some cases, experienced utility is the criterion by which a decision should be assessed, because an individual who pays different amounts to achieve the same gain of experienced utility is making a mistake (Kahneman, 2011, pp. 368, 369).

These conclusions lead to the question how to measure experienced utility? How should we answer questions such as: “How much pain did Helen suffer during the medical procedure?” or “How much enjoyment did she get from her 20 minutes on the beach?” (Kahneman, 2011, p. 369). Many economists tried to measure utility, but one of them created a coherent theory which was used by Kahneman to measure experienced utility. It was British economist Francis Edgeworth who in the 19th century proposed the idea of a ‘hedonometer.’ It was an imaginary instrument analogous to the devices used in weather-recording stations. But instead of measuring temperature, it measures the level of pleasure or pain that individual experiences at any moment (see (Edgeworth, 1879)). Experienced utility varies, as daily temperature does, and the results were presented as a function of time. Therefore, the answer to the question how much pain or pleasure Helen experienced during her medical procedure or vacation would be the “area under the curve” (Kahneman, 2011, p. 369). In Edgeworth’s conception time plays an essential role. If Helen stays on the beach for 60 minutes instead of 20 and her enjoyment remains as intense, then the total experienced utility of that episode triples.

Kahneman and Redelmerier designed the experiment to see how hedonometer would work in reality. Figure 1 shows profiles of the experiences of two patients undergoing a painful colonoscopy.⁸ The patients were prompted every 60 seconds to assess the level of pain they experienced at the moment. The data shown are on a scale where zero is ‘no pain at all’ and 10 is ‘intolerable pain.’

If people were asked which of these patients suffered more, everyone would answer that patient B had the worse time. Patient B spent at least as much time as patient A at any level of pain, so the ‘area under the curve’ is evidently larger

⁷ The fact that people look at relative not absolute outputs is also one of the conclusions from ‘prospect theory’ (Kahneman, & Tversky, 1979).

⁸ This procedure is now routinely administered with an anaesthetic as well as an amnesic drug, but these drugs were not as widespread when Kahneman collected his data.

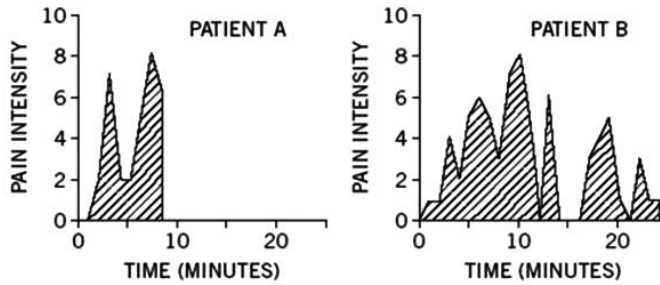


Figure 1. Pain intensity

Source: (Kahneman, 2011, p. 370).

for B than for A. Of course, the most important factor here is that patient B's procedure lasted much longer. However, this simple answer was not in accordance with the results of the participants' surveys. After the procedure, all participants were asked to rate "the total amount of pain" they had experienced during the procedure. The wording was designed to encourage them to think of the whole of the pain they had reported, reproducing the hedonometer. Surprisingly, the patients did not perceive pain in the hedonometer scale. The statistical analysis revealed two findings which illustrate a pattern that was observed in the other experiments:

Peak-end rule: The global retrospective rating was well predicted by the average of the level of pain reported at the worst moment of the experience and at its end.

Duration neglect: The duration of the procedure had no effect whatsoever on the ratings of total pain. (Kahneman, 2011, p. 370)

As it turned out, patient B perceived colonoscopy as less painful than patient A. It was possible, because people do not perceive pain in the hedonometer scale, but they use 'peak-end rule'. The worst ratio (8) was the same for both patients, but the last rating before the end of the procedure was 7 for patient A and only 1 for patient B. The peak-end average was therefore 7.5 $[(8 + 7)/2]$ for patient A and 4.5 for patient B $[(8 + 1)/2]$. In the result, patient A retained a much worse memory of the episode than patient B. This experiment shows that two systematically different measures of experienced utility exist – the hedonometer total and the retrospective assessment. The hedonometer total is computed by an observer from an individual's report of the experience of moments. The equal weight is assigned to all moments: three minutes of pain at level 8 is three times as bad as one minute at the same level of pain. On the other hand, the findings of this experiment point out that the retrospective assessments are insensitive to duration. What matters are two singular moments – the peak and the end. They indicate how good or bad are people's memories of something.

These findings mean that people may have two different selves which perceive reality differently. The experiencing self answers the question: “Does it hurt now?” The remembering self answers the question: “How was it, on the whole?” Depending on the particular self we will get a different answer about an experience. Kahneman carried the experiment to show the conflict between experiencing self and remembering self. The participants were asked to hold their hands in painfully cold water and afterward they were offered a warm towel. The experiment goes like this:

Each participant endured two cold-hand episodes: The short episode consisted of 60 seconds of immersion in water at 14° Celsius, which is experienced as painfully cold, but not intolerable. At the end of the 60 seconds, the experimenter instructed the participant to remove his hand from the water and offered a warm towel.

The long episode lasted 90 seconds. Its first 60 seconds were identical to the short episode. The experimenter said nothing at all at the end of the 60 seconds. Instead he opened a valve that allowed slightly warmer water to flow into the tub. During the additional 30 seconds, the temperature of the water rose by roughly 1°, just enough for most subjects to detect a slight decrease in the intensity of pain. (Kahneman, 2011, pp. 372, 373)

After the second trial, the participants were given a choice about the third (last) trial. They could decide which of the previous trials they want to repeat. This choice causes not only a conflict between experiencing and remembering self, but also between experienced utility and decision utility. From the perspective of the experiencing self, the long trial is clearly worse. The situation is different when we refer to the remembering self. Due to the peak-end rule, people will have a worse memory of the short than of the long trial. Moreover, the duration neglect indicates that the difference between 90 second and 60 seconds will not matter at all.

Eighty percent of participants behaved in accordance with remembering self. They chose to repeat the longer episode, which means that they willingly agreed to suffer 30 seconds of needless pain. What is crucial here is that these people were not masochists and they did not intentionally choose to suffer more. They just made a mistake. If we had asked them “Would you prefer a 90-second immersion or only the first part of it?” they would certainly have selected the short option (Kahneman, 2011, p. 373). Kahneman did not use these words, because he wanted to get natural responses. In the experiment, people chose to repeat the trial about which they had a better (less painful) memory. Their decision was governed by a simple rule of intuitive choice – pick the option which you like the most or dislike the least. In the end, it is a memory which determined what people would choose. Kahneman (2011, p. 375) explains why the implications of the above-mentioned experiment are so vital for neoclassical economics:

Decisions that do not produce the best possible experience and erroneous forecasts of future feelings – both are bad news for believers in the rationality of choice. The cold-hand study showed that we cannot fully trust our preferences to reflect our interests, even if they are based on personal experience, and even if the memory of that experience was laid down within the last quarter of an hour! Tastes and decisions are shaped by memories, and the memories can be wrong. The evidence presents a profound challenge to the idea that humans have consistent preferences and know how to maximize them, a cornerstone of the rational-agent model. An inconsistency is built into the design of our minds. We have strong preferences about the duration of our experiences of pain and pleasure. We want pain to be brief and pleasure to last. But our memory, a function of System 1, has evolved to represent the most intense moment of an episode of pain or pleasure (the peak) and the feelings when the episode was at its end. A memory that neglects duration will not serve our preference for long pleasure and short pains.

Moreover, the existence of two distinct selves such as remember self and experiencing self, is one of the arguments for multiple self conception which will be scrutinized more in Chapter 4. The knowledge that people experience things differently and that experience depends on a self to which we refer imply that we need to discuss each self separately.

The next important problem with preferences is that people do not know what preferences they will have in the future. People are not good, and they are not perfect for sure, at foreseeing how much utility they will gain from future consumption. One of the most significant psychologists of our time dedicated the whole book to show that people have enormous problems to assess the future flow of happiness ((Gilbert, 2007), see also (Loewenstein, & Schkade, 1999)). The most significant mechanism that makes people unable to know what they will want in the future was presented earlier in the case of present bias. Now, the two supplement phenomena will be analysed. First of them is called ‘focus illusion’ and it was introduced by Schkade and Kahneman (1998). It can be summarized by the moral and warning “Nothing that you focus on will make as much difference as you think” (Schkade, & Kahneman 1998, p. 353). It means that individuals tend to exaggerate the importance of any aspect of their lives when they focus their attention on it. Schkade and Kahneman (1998) carried out the experiment in which many students from Midwest and Southern California rated their overall life satisfaction as well as various aspects of life, for either themselves or someone similar to them in one of these two regions. It can be concluded that this self-reported life satisfaction was the same in these two regions but these participants who rated another group expected Californians to be more satisfied than Midwesterners. It means that people intuitively (wrongly) assume that life in California must be better, because they focus their attention on one factor (weather). It is wrong, because people forget that weather is not

the only thing that influences our life satisfaction. What is also important is job, family, social connections and many other factors.

Another problem with preferences over time is when people make a decision in a 'hot state', e.g., buying automobiles when the dealer makes them excited, getting married in the heat of passion, or committing suicide in the depth of depression (Loewenstein, O'Donoghue, & Rabin, 2003). 'Hot state' results from present bias. What is more, it is difficult for people to realize that the state of arousal will pass quickly and their decisions will be seen as irrational.

As we can see, the revealed preferences theory is not always a good indicator of what gives people happiness. People often do not know what they want due to various cognitive biases. Therefore, choice alone does not indicate what makes people happy. Moreover, it does not always represent human preferences correctly. This is why economics of happiness has a role in analysing human happiness.

It is a cliché to say that people want to be happy and that happiness has a crucial role in people's life. However, it ceased to be so obvious when one looks at value-free economics and its reluctance towards happiness. Therefore, some sentences are necessary if we want to understand the importance of happiness. In the end, happiness is treated by economics of happiness as the ultimate goal which should be maximized. Measuring happiness is only a secondary issue which is intended to help people to establish what gives them happiness.

The importance of happiness was indicated more than two millennia ago by the famous sentence which is attributed to Aristotle "happiness is the meaning and purpose of life, the whole aim and end of human existence." For him happiness was an ultimate goal to which other things were only the means. Aristotle ascribed to wealth much less importance: "The life of money-making is one undertaken under compulsion, and wealth is evidently not the good we are seeking; for it is merely useful and for the sake of something else" (Aristotle, 2000, I, 1096a. 5). The importance of searching for happiness was generally recognized in ancient moral philosophy. However, there was not any consensus about how to achieve happiness - in those times there was a lot of various schools with different approaches towards happiness (e.g., Cynicism, Stoicism, Epicureanism, Aristotelianism). Despite their disagreements about ways to achieve happiness, they agreed that happiness was a crucial part of human life. Aristotle (2000, I, 4, 1095a) expressed this belief:

Most people, I should think, agree about what [...] is called [the highest good], since both the masses and sophisticated people call it happiness, understanding being happy as equivalent to living well and acting well. They disagree about substantive conceptions of happiness.

This perception of happiness as the ultimate end in life is absolutely different from the economics point of view in which happiness is interpreted within the framework of a means-ends relationship (value-free economics).

3.4. Economics of happiness – general insights

Economics of happiness has produced numerous interesting insights into the determinants and the nature of happiness. Frey, one of the most prominent supporters of this new approach, presented some of the major results:

- Most people, during most periods, in most countries, are satisfied with their lives.
- Economic conditions – income, employment, price stability, fair income distribution – are important determinants of happiness.
- Non-material aspects – family, friendships, other social ties – matter greatly for happiness.
- People tend to adjust to their basic level of happiness after positive and negative life events, but the speed and the extent of such adjustment differ depending on whether it relates to income, employment or other areas.
- People are status seekers and always compare themselves to others.
- Marriage makes people happy – but not for long.
- Children lower parents' life satisfaction, but make them happy once they leave the household.
- Extensive television viewing makes active people less happy.
- Helping others by volunteering and by giving financial support increases happiness.
- People make systematic errors about their happiness with respect to both the past and the future (misprediction); they are subject to weakness of will.
- Procedural utility matters to happiness above and beyond outcome utility.⁹
- Culture has little effect on the marginal effect of the various determinants of happiness.
- Political institutions – in particular, opportunities for citizens to participate via democracy and federalism – are significant determinants of life satisfaction. (Frey, 2010, p. 154)

From the perspective of value-free economics, two theoretical points made by economics of happiness are much more significant (Frey, 2010, p. 200). First of them states that it is possible to use the measures of subjective well-being as

⁹ Procedural utility means that people do not only value actual outcomes, i.e., the what, but also the conditions and processes that lead to these outcomes (Frey, Benz, & Stutzer, 2004). We can find analogy here with Aristotle's virtue ethics. In general, economics of happiness frequently refers to Aristotle, see (Angner, 2010; Bruni, Comim, & Pugno, 2008; Bruni, & Porta, 2007).

a (reasonably) good proxy for the theoretical concept of utility as preference satisfaction. Earlier a few methods of measuring subjective well-being were presented. Each of them has its problems and can be easily criticized. However, these methods are improving all the time and now they seem developed enough to use them as a guide for economic theory and policy. The second claim is that the measurement of utility is necessary to cope with some issues which economics has to face. Utility should be measured in the cases when revealed preferences do not faithfully reflect the individual's utility. Due to cognitive biases, people make systematic errors in the decision making, which are especially visible when people make choices over time.

These two claims that utility should be measured and can be measured stand in opposition to the foundation of value-free economics. This is why economics of happiness is perceived as a counter-revolution to standard microeconomics (Frey, 2010, p. 200). However, due to the weaknesses of subjective well-being, there is a need for a different approach to happiness. Objective well-being approaches criticize subjective well-being and they try to introduce some ultimate goals that can give people happiness.¹⁰ Subjective well-being seems as a step in the right direction in terms of explanatory and predictive power, but it is not enough, because of the problems that this concept generates.

3.5. Objective well-being approaches

Before objective well-being approaches are going to be examined, it is necessary to highlight the problems with the ambiguity of the term happiness. This ambiguity leads to the situation in which it is difficult to distinguish between subjective (economics of happiness) and objective approaches towards well-being (e.g., the capability approach). Happiness can mean literally everything. Easterlin (2001, p. 465) illustrates this problem: "I use the terms happiness, subjective well-being, satisfaction, utility, wellbeing, and welfare interchangeably." As was indicated earlier in this book, economists do not even like the question 'what is happiness?' For many of them, it is an elusive term which cannot be scientifically analysed. Unlike economists, psychologists deal with happiness, but they also acknowledge the ambiguity of the term happiness:

¹⁰ The different approaches to happiness (1 - subjective well-being, 2 - revealed preferences, 3 - human flourishing) presented in this chapter are similar to the approaches indicated by Parfit: "What would be best for someone, or would be most in this person's interests, or would make this person's life go, for him, as well as possible? Answers to this question I call theories about self-interest. There are three kinds of theory. On Hedonistic Theories, what would be best for someone is what would make his life happiest. On Desire-Fulfilment Theories, what would be best for someone is what, throughout his life, would best fulfil his desires. On Objective List Theories, certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things" (Parfit, 1984, p. 493).

A widely presumed component of the good life is happiness. Unfortunately, the nature of happiness has not been defined in a uniform way. Happiness can mean pleasure, life satisfaction, positive emotions, a meaningful life, or a feeling of contentment, among other concepts. (Diener, Scollon, & Lucas, 2004, p. 188)

The ambiguity of happiness is a problem when it comes to defining well-being. Therefore, it is necessary to differentiate between the hedonic and eudaimonic perspectives.

Hedonic approach is equivalent to economics of happiness. It is based on utilitarian tradition (Bentham),¹¹ where pleasure and pain are people's sovereign masters and they guide the individual's decision-making. Nowadays, hedonic approach has a stronger foundation than utilitarianism, because pleasures and pains (SWB) can be measured objectively at least this is what its supporters claim (e.g., (Kahneman et al., 1999)). In his book, Kahneman declared the existence of a new field of psychology. Its title, *Well-being: The foundations of hedonic psychology*, evidently suggests that, within this paradigm, the terms 'hedonism' and 'well-being' are equivalent. In this approach, well-being is perceived as vaguely as the term 'happiness.' Deci and Ryan (2001, p. 144) indicate this broadness: "hedonism, as a view of well-being, has thus been expressed in many forms and has varied from a relatively narrow focus on bodily pleasures to a broad focus on appetites and self-interests". It is essential to understand that in hedonic approach 'higher' feelings such as altruism, love or sacrifice are assessed, because in general opinion hedonism is simplified to physical pleasures. This criticism started with Epicure and his system which was named "the philosophy of swine" (Weijers, 2011). The opponents did not understand that pleasures do not necessarily need to come from physical, 'lowest' activities. The same criticism fell on Bentham who writes that "prejudice apart, the game of push-pin is of equal value with the arts and sciences of music and poetry" (Bentham, 1825, p. 206). In his opinion, both activities are equally good if they give the same amount of utility. However, it does not mean that people cannot get more pleasure from reading Joyce's *Ulysses* than from watching a soap opera. What is important in hedonic approach are mental states (subjective well-being), but what gives us happiness is relative.

The alternative variant of hedonic approach has been developed in the Aristotelian tradition and it is labeled as eudaimonic approach. In this approach, decision making is governed by the assessment of what constitutes a good life. Nowadays, *eudaimonism* and *hedonism* are the two major approaches in the field of happiness research in economics. In eudaimonic approach well-being is something more than only hedonic or subjective happiness. Ryan and Deci (2001, p. 145) point

¹¹ One of the main supporter of hedonic approach is Kahneman who wrote the article which has a symptomatic title: *Back to Bentham? Explorations of experienced utility* (Kahneman, Wakker, & Sarin, 1997).

out that “despite the currency of the hedonic view, many philosophers, religious masters, and visionaries, from both the East and West, have denigrated happiness *per se* as a principal criterion of well-being”. In eudaimonic approach well-being is understood as the realization of human potential which is perceived as one’s *daimon* – true nature. *Eudaimonia* cannot be identified with pleasure. Bruni points out that “neo-Aristotelian philosophers in the Anglo-Saxon world preferred to translate *eudaimonia* as ‘human flourishing’ rather than happiness, because in common language today happiness also indicates momentary euphoria, carefree content, a pleasurable sensation or *tout court* pleasure” (Bruni, 2008, p. 124). Philosopher Martha Nussbaum defines *eudaimonia* as general state, not as the output of some activities: “Happiness is something like flourishing human living, a kind of living that is active, inclusive of all that has intrinsic value, and complete, meaning lacking in nothing that would make it richer or better” (Nussbaum, 2005, p. 171). Therefore, in objective well-being approaches (human flourishing, *eudaimonia*) happiness is only a part of well-being. It is still a significant part, but happiness is not the only goal which is worth achieving.

Sometimes, even, pursuing happiness can be a bad idea. *Eudaimonia* cannot be achieved only through instrumental means. It is the indirect result of actions that are carried out because of their intrinsic value. If we treat happiness as something that can be achieved, then the result can be opposite to our intentions. This phenomenon is called ‘paradox of happiness’ and J.S. Mill was one of the thinkers who described it. He writes in his autobiography:

But I now thought that this end [one’s happiness] was only to be attained by not making it the direct end. Those only are happy (I thought) who have their minds fixed on some object other than their own happiness [...]. Aiming thus at something else, they find happiness along the way [...]. Ask yourself whether you are happy, and you cease to be so. (Mill, 1909, p. 94)

An illustration of this paradox was offered by Schooler, who studied the “costs of trying to have a good time” on New Year’s Eve (Schooler, Ariely, & Loewenstein, 2003, p. 59). Before this big event researchers sent an e-mail which contained a questionnaire to 475 people. The questions referred to things such as how large celebration they were planning, how much they expected to enjoy it, and how much money and time they were expecting to spend on it. After the event, people were asked the same questions with regard to their actual experiences. The researchers found that those people who expected a great party were more likely to be disappointed than those who expected only a small celebration or none at all. The difference between expected and experienced enjoyment was negatively correlated with people’s anticipation and with the time they expected to spend on preparations. It means that people who really want and try to have a good time ended more unhappy than people who did not try so much.

In objective well-being approaches happiness is not everything that matters. There are two reasons why the advocates of well-being approaches are not fully convinced to hedonic approach. Firstly, they claim that an individuals' preferences are too malleable. Social environment strongly influences individuals' preferences and this made them unreliable as the indicator of well-being. Secondly, the concept of well-being contains many aspects and preferences do not provide satisfactory information to assess some of these aspects (e.g., freedom, autonomy) (Sugden, 2008a).

The most important psychological mechanism that makes an individual's preferences malleable is a process called 'hedonic adaptation' or 'hedonic treadmill.' The seminal paper which described this phenomenon was published by Brickman, Coates, and Janoff-Bulman (1978). The researchers interviewed people who had won the lottery over the last year (average \$ 479,545) and the control group. The researchers had not found any difference in reported happiness between the winners and non-winners. A more striking example of hedonic adaptation is that people who won big lotteries after a year were at similar level of happiness as people who had an accident and had to use a wheelchair till the end of their lives (Brickman et al., 1978). Hedonic adaptation demonstrates that people adjust to a new situation. At the beginning a new situation influences people's subjective well-being (winning a lottery - happy; being on a wheelchair - unhappy) but with time people adapt to a situation and they return to the previous level of happiness. One of the reasons for hedonic adaptation is that it can protect people from situations which diametrically diminish their well-being (e.g., death of loved ones, accident). It would be tough for people to function well in despair and this is why with time we feel better. Hedonic adaptation is expressed in one of the most known proverbs "time is a great healer." The phenomenon of hedonic adaptation (Brickman et al., 1978) immediately gained popularity and acceptance and has even been called the first principle of happiness research (Myers, 1992). Although hedonic adaptation greatly influences people's behaviour, it does not mean that people can adjust to everything. The psychological studies on hedonic adaptation have also shown that some things are more prone to habituation than others (Frederick, & Loewenstein, 1999, p. 311). People do not get used to the persistent stress of commuting, pollution, noise, or loneliness.

Directly connected with hedonic adaptation is 'set-point theory.' In this theory, the major life events like the birth of a child, the death of a spouse or a debilitating disease affect people's happiness only temporarily. After some time, however, a person's happiness level returns to a default level which is determined by personality and genetic traits. This theory explains why some people who had a hard life are happy and peaceful when people who have achieved everything both in their careers and family lives are unhappy and dissatisfied.

Hedonic adaptation is interchangeably called ‘hedonic treadmill.’ Here, the focus goes to a situation which improves people’s life. It is called hedonic treadmill, because it is impossible for people to be satisfied with what they have, because they will always want more. For example, if one buys a new car for a few weeks one will be very happy but after a few months this car will not be a new exciting car but a regular, boring one. If one gets a promotion at work at first, one will feel very happy but after a few months a new position will be something normal. This mechanism is very powerful, because people easily adjusted to a new, better situation and treat it as something normal. This situation is connected with aspiration level theory where a person’s happiness is based not only on one’s attainment but also on one’s aspiration level. Therefore, if one gets a promotion, for a while one is happy, but after some time one seeks for a next better position, and one’s current situation in comparison does not look so good. This is why people with higher aspiration are associated with lower happiness (Frey, & Stutzer, 2002b; Stutzer, 2004). Thus, the paradox of “happy peasants and frustrated achievers” which was introduced by Graham (2009, p. 19). One’s aspirations came mainly from comparisons with others. Marx famously writes about the relativity of comparisons:

A house may be large or small; as long as the neighboring houses are likewise small, it satisfies all social requirement for a residence. But let there arise next to the little house a palace, and the little house shrinks to a hut. (Marx, 2006, p. 33)

The influence of comparisons is so significant that they can even make people irrational. The well-known example of this mechanism is the case of Olympic winners. When psychologists were watching the Olympic winners’ photos, one thing immediately draws the eye. Silver medallists were often visibly less happy than contestants who won the bronze medal. The research carried on by Medvec found that people who won the Olympic bronze medals reported being happier than silver medallists (Medvec, Madey, & Gilovich, 1995). The reason for this is ‘aspiration theory.’ Silver medallists imagine themselves as the winners of the gold medal, and they feel disappointed, because they could be the best in the world. On the other hand, bronze medallists were very happy, because they imagined themselves in the 4th place which is the worst place for the Olympic contestants (so close to medal). From this perspective, the third place looks really attractive.

The next problem with preferences and their malleability is connected with the ‘experience-stretching hypothesis.’ This hypothesis can answer to the problem which was presented earlier in the context of the relativity of subjective well-being. The question was if we can tell that Mark who declared to feel eight on the 10 points scale is happier than Tom who declared seven. The experience-stretching hypothesis convinces us how important is the context within which people make their decisions. For example, a scuba diver who experienced Great

Barrier Reef would say about the excited beginner who has immersed in the lake for the first, that “he is excited because he does not know what is really beautiful.” It is the case, because their experiences are different. A beginner is excited, because for him a submersion in the lake is 10/10. For the experienced scuba diver, Great Barrier Reef was 10/10, and he does not find anything exciting in short immersion in a lake (1/10) (Michoń, 2010, pp. 77, 78). This means that it is very problematic to compare their level of subjective well-being, because the scale which is used depends on personal experience. Gilbert writes that it is impossible to decide which scale of subjective well-being is better (being a happy beginner or more experienced):

So which hypothesis is correct? We can't say. What we can say is that all claims of happiness are claims from someone's point of view—from the perspective of a single human being whose unique collection of past experiences serves as a context, a lens, a background for her evaluation of her current experience. As much as the scientist might wish for it, there isn't a view from nowhere. Once we have an experience, we are thereafter unable to see the world as we did before. Our innocence is lost and we cannot go home again. (Gilbert, 2007, p. 57)

These mechanisms (hedonic adaptation/treadmill, set point theory, aspiration theory, experience stretching) demonstrate that preferences and subjective well-being are malleable and that they are not good indicators of well-being. These mechanisms also explain ‘Easterlin paradox.’ People from very poor countries have a similar level of subjective well-being to people from highly developed countries. The main reason for this situation is that people from wealthy countries are not so much satisfied with their lives because of hedonic adaptation and higher aspirations, whereas people from poor countries do not know how much better their life could be (experience stretching).

Sen has been dealing with the underdeveloped countries for most of his scientific career. He used the example of people who are poor and deprived of freedom, but they consider themselves happy. He wrote:

Consider a very deprived person who is poor, exploited, overworked and ill, but who has been made satisfied with his lot by social conditioning (t[h]rough, say, religion, political propaganda, or cultural pressure). Can we possibly believe that he is doing well just because is happy and satisfied? Can the living standard of a person be high if the life that he or she leads is full of deprivation? The standard of life cannot be so detached from the nature of the life the person leads. (Sen, 1985b, p. 12)

By this example, Sen wants to point out that self-reported happiness is not sufficient to be the only indicator of well-being and it must be based on a more objective basis. Sen originated such a movement which is called the ‘capability approach,’ but this topic will be scrutinized in Chapter 4. Let us get back to the example with deprived people. The above-mentioned illustration is not as extreme as it looks. For example, many women from the Arabic countries

considered themselves as happy even though they cannot vote, drive a car, go outside the house without a male member of the family and they must be absolutely obedient to the husband (see (Nussbaum, 2000)). Of course, a cultural or religion pressure can be considered as a specific case. However, the point made by Sen is much broader. He uses the hedonic adaptation to explain why some people adjust to awful lives and why they perceived themselves as happy or at least not unhappy people. As Sen (1985a, pp. 14, 15) puts it:

A person who is ill-fed, undernourished, unsheltered and ill can still be high up in the scale of happiness or desire-fulfilment if he or she has learned to have 'realistic' desires and to take pleasure in small mercies. The destitute thrown into beggary, the vulnerable landless labourer precariously surviving at the edge of subsistence, the overworked domestic servant working round the clock, the subdued and subjugated housewife reconciled to her role and fate, all tend to come to terms with their respective predicaments. The deprivations are suppressed and muffled in the scale of utilities (reflected by desire-fulfilment and happiness) by the necessity of endurance in uneventful survival.

Sen uses hedonic adaptation to criticize both subjective well-being (economics of happiness) and desire-fulfilment theory (neoclassical economics). For him being happy or being able to satisfy one's preferences does not mean that these people lead a good life. Happiness and preferences are too malleable, because people can adjust even to the dire life condition. However, if one wants to state that the lives of people in the above-mentioned examples are not good, one needs to have some goals and values which are independent from subjective well-being and revealed preferences. The attempt to find these goals will be presented in Chapter 6.

We have just established that preferences are malleable but it is not the end of the problems. The proponents of objective well-being argue that well-being contains many aspects which cannot be adequately assessed by preferences or happiness (Bruni et al., 2008). The most critical aspect is freedom, because often it stands in opposition to happiness and people need to decide which of them they want to have. Of course, everyone wants to be happy. However, people also perceive liberty as an intrinsic value (deontological perspective). This fact is demonstrated in the Declaration of Independence where it is stated that "we hold these Truths to be self-evident, that all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty, and the Pursuit of Happiness" (*The Declaration of Independence*, 1776).

Now, let us move into a few examples which demonstrate the dichotomy between happiness and freedom. This dichotomy is the reason for the differences between hedonic and capability approach. The example created by Griffin (1986, 8 ff.), which was later expanded by Barrotta (2008, pp. 8, 9), is presented to highlight

that the conception of happiness in hedonic approach is too narrow and it does not contain different matters that people have reason to value. The example is about Freud who was very sick at the end of his life. He refused to take any drugs, because he wanted to think clearly even though it meant to live in very painful conditions. In hedonic approach, thinking freely is nothing more than a possible source of happiness. However, it seems plausible to presume that neuroscientists would figure out that not taking drugs causes a decrease in Freud's happiness. The significant question that should be answered here is: if Freud made a wrong choice to refuse to take the drugs since his decision lowers his level of happiness? Barrotta (2008, p. 151) emphasizes that this problem is not empirical:

I really do not know what neuroscientists would say about my example, though I find it plausible that they would measure a decrease in pleasurable feelings. I am only assuming that it is conceptually *possible* that a subject deliberately makes a decision that reduces his or her hedonic payoff, as measured by neuroscience.

However, this conceptual possibility is the puzzle for hedonism. If people had knowledge, they should choose to be happier than less happy. Therefore, if we know that the decision of Freud diminishes his happiness, then we are forced to conclude that Freud made the wrong choice. However, Freud to show that he is not wrong could reply:

You should not consider only the level of my pleasurable feelings, but also my concern for such feelings. Namely, you should also consider how much I value them and how important they are in my life. As a matter of fact, I did not make the wrong choice. Rather you employ too narrow a definition of happiness. (Barrotta, 2008, p. 152)

Barrotta argues that Freud is right. Thinking freely is not solely a source of happiness. If it were, then we would be able to compare its consequences with the consequences of the decision to take drugs. Thinking clearly is something independent of happiness and it has value *per se* which can be the goal as important as happiness. In this way, consequences (happiness) are not the only important factor in the process of choice. Individuals have freedom to make their own evaluations about their feelings of their mental states. It means that they can decide what a happy life is. If one wants to make sense of Freud's answer, one needs to include 'thinking freely' in Freud's idea of a happy life. Therefore, as Barrotta (2008, p. 152) points out "happiness is an intrinsically pluralistic concept". There is no single objective idea of utility or happiness which can be assessed by scientists. People have the freedom to choose things that they value even if these things do not give people the most 'objective' happiness. The ability to assess one's mental states and not being a slave of hedonic conception of happiness is one of the most critical issues in the capability approach which will be analyzed in Chapter 6.

The next example which shows that people can choose some other things than pleasure was presented by Nozick (1974, pp. 42–45). He created a thought experiment which is known as ‘The Experience Machine’ where the person’s brain is connected to some machine. As a result, the subjects of simulation fall asleep and they have a virtual and extremely pleasant experience. Moreover, the subjects of the simulation would think that these pleasant experiences are a reality. Many people do not want this kind of life even if it would be delightful.¹² The rejection of an extremely pleasant life is an argument against hedonism, because it demonstrates that people value other things more than pleasure. The reluctance to use only utilitarian measurement justifies the acknowledgment of a possibility that there are other values than pleasantness or unpleasantness.

The next example also indicates that happiness is not the only important thing in life. We can go back to the criticism of utilitarianism to see that living in fool’s paradise is not the best option. One of the arguments against Bentham and his idea of happiness was that the very happy oyster which lives a very long time could have a better life than an average human. J.S. Mill answered this criticism towards utilitarianism (swine philosophy) when he wrote:

It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are a different opinion, it is because they only know their own side of the question.
(Mill, 1863, p. 14)

Mill argues that there are better (intellectual) and worse (connected with physical sensation) pleasures and if a person has experienced both pleasures, she will choose the better. In the next chapter, it will be argued that this argument is not a utilitarian one. Now, this example is used to demonstrate that people want some other things than simple pleasures, and that the fool’s paradise is not a good option if one knows the very nature of it. This is why Mill in some way forestalled Gilbert who wrote about experience stretching and where he argued that “once we have an experience, we are thereafter unable to see the world as we did before. Our innocence is lost and we cannot go home again” (Gilbert, 2007, p. 57).

The dichotomy between happiness and freedom is a topic which is very interesting for social thinkers who wrote many books about the dystopias in which people are happy but for the price of freedom. Probably the most known book about this topic is *New brave world* written by Aldous Huxley (2006). People

¹² Recently, however, some researchers challenged Nozick’s assumption that most people choose reality not an artificial and blissful life. One of them is De Brigard who carried the empirical studies which show that “people’s intuitions about the experience machine in fact are highly divergent” (De Brigard, 2010, p. 55) and some people choose happy life over reality.

in this book besotted themselves with happiness pills called ‘soma.’ These pills desensitize people to the pain of the existence (*Weltschmerz*) but transform people to the ‘happy pigs.’ The conclusion from this book is presented as a warning to artificial happiness which causes that people do not look like people anymore.

The last problem with the hedonic approach concerns reality. A lot of researchers who investigate happiness believe that it is possible to measure happiness objectively (e.g., (Harris, 2011; Kahneman et al., 1999; Layard, 2003, 2005). This belief is not a problem itself, but its implications are. Many happiness researchers think that state, in various forms and to different degrees, should decide on behalf of people (for their own good). This approach comes as no surprise when one has in mind that in hedonic approach happiness is the ultimate goal. Therefore, subjective well-being is considered to be the only criterion of the progress of society and other values like freedom are only elements of well-being. In the case of Freud, the supporters of hedonic approach (Layard, 2003, 2005)¹³ would argue that if ‘thinking freely’ diminish happiness, it is irrational for Freud to choose this option. If happiness is an objective goal for everyone, then it is not so peculiar to propose the idea that the state should coerce Freud to take the drugs (for his own good). These conclusions are disturbing and they were acknowledged by the happiness approach supporters. Frey and Stutzer (2010, p. 567) argue that the state orientation on a happiness index reduces citizens to “metric stations” and people are not treated as autonomous, sovereign agents. In Chapter 7, however, the author will check if it is possible to have both happiness and freedom on the example of libertarian paternalism.

3.6. Conclusions

In this chapter, the most significant insights from economics of happiness were analysed to check whether the axiomatic assumption of value-free economics (people do what is the best for them) reflects reality and serve people in the context of happiness. The conclusion is that this assumption is not correct and people frequently do not make the best choices (Bruni et al., 2008; Bruni, & Porta, 2005, 2007; Frey, 2010, 2018; Frey, & Stutzer, 2002b; Kahneman, et. al., 1999). It is the result of many cognitive biases that people have which were indicated by behavioural economics (Frank, 1988; Kahneman, 2003a, 2003b, 2011; Kahneman, & Tversky, 1979; Thaler, 1980, 2000, 2015; Loewenstein, 1996; Loewenstein, & Elster, 1992; Loewenstein et al., 2003; O’Donoghue, & Rabin, 1999, 2001; Rabin, 1998; Simon, 1957). Both economics of happiness and behavioural economics indicate the importance of the environment within which individuals make

¹³ Layard (2005, p. 113) writes that “we naturally look for one ultimate goal that enables us to judge other goals by how they contribute to it. Happiness is that ultimate goal, because, unlike all other goals, it is self-evidently good.”

their choices. Thanks to that, we learn that one of the essential assumptions of revealed preference theory – consistency – cannot be fulfilled and that people always choose from some “menu dependence” (Sen, 1997, p. 752).

Economics of happiness, even for being able to indicate that people fail to maximize their utility/happiness, needs some measurement of happiness that is not axiomatic and where it is not assumed that what people do is the best for them. The primary measurement that economics of happiness uses is SWB (subjective well-being). It is a method of measurement which is independent of the revealed preference theory which means that maximization of happiness is not identified with revealed choices as it takes place in value-free economics.

However, subjective well-being measure has its own problems which in some way resemble the obstacles that were encountered by value-free economics many years ago. The most significant complication is the relativity of happiness and utility. Relativity is an inherent feature of SWB (Alexandrova, 2005; Angner, 2010). On the one hand, it can be perceived as an advantage, because who be better to judge how people feel than people themselves. On the other hand, people are not experts in predicting what will give them happiness and frequently they do not choose what is best for them.¹⁴

Due to its relativity, SWB measures face very similar problems as value-free economics. Happiness in hedonistic approaches is also the ‘black box’ that contains all categories such as happiness, freedom, love, justice, etc. Therefore, economics of happiness does not distinguish other values from happiness. The next problem that concerns relativity in hedonistic approaches is the psychological fact that people’s preferences are malleable. Thus, people can adjust even to horrible conditions (Sen, 1985a, pp. 14, 15).

Because of these problems, it is necessary to examine the approaches that indicate objective goals/values and are focused on human flourishing. Objectivity here does not mean objectivity in the sense of the absolute Truth but it is rather an indication of goals which are independent of individuals’ subjective perception of happiness and from individuals’ choices (revealed preferences). The eudaimonic approach was presented shortly to point out the main differences between objective approaches and hedonistic approaches. In Chapter 6, one of the objective approaches (capability approach) is analysed in depth to decide whether this conception can resolve the problems of hedonistic approaches (1 – malleability of preferences, 2 – other values than happiness).

It should be noticed that objective approaches towards happiness as well as value-free economics in the 1930s tried to react to the relativity of both happiness

¹⁴ Sometimes people know that they do not choose the best option, but do this nonetheless. To explain this paradox, the phenomenon of weakness of will is analysed in the next chapter.

and utility. Value-free economists abandoned cardinal utility and created the axiomatic approach in which everything is logically consistent. The proponents of objective approaches reacted differently. They have been trying to find objective goals independent of people's choice.

To sum up. Despite criticism, economics of happiness (hedonic approach) is a step in the right direction for economics in terms of explanatory and predictive power. It gives answers to the questions that value-free economists were not even willing to ask. Economists need to analyse happiness, measure it and investigate which things give people happiness. From this perspective revealed preference theory has limited cognitive value. It is in particular in the case of bounded rationality which demonstrates that people do not make the best decisions. However, subjective well-being approach has some problems which can only be resolved by the introduction of goals and values independent of happiness/utility and revealed preferences. This is crucial if one looks at the choices over time and the conception of multiple self. In the next chapter, the author will check whether it is possible to compare utility over time. If the answer is no, then we will need to choose normatively between different selves ("short-time human" and "long-time human" (Ostapiuk, 2019b)). In order to choose between selves, one needs to know which self is superior and for this purpose we need to justify the existence of values which are independent of revealed preferences. Some normative approaches are examined to check what people should do to achieve both happiness and freedom (multiple self, capability approach, and libertarian paternalism).

Chapter 4

The conception of multiple self. Reason versus emotions

The only way in which the economist can keep his studies from duplicating the psychologist's work is by taking his psychology from those who have specialized in that field. [...] The economist may attempt to ignore psychology, but it is a sheer impossibility for him to ignore human nature, for his science is a science of human behavior

(Clark, 1918, p. 4)

4.1. Introduction

This chapter has a number of goals. The first of them is to demonstrate that neoclassical/value-free economics has a simplified image of the individual. The philosophical discussion on personal identity is presented to indicate this. After this discussion it will be possible to introduce and argue for the conception of multiple self. The main focus is on choices over time which cannot be compared by utility measurements. Because of hyperbolic discounting, we need to decide which self should be in control. In the end, the phenomenon of weakness of will is analysed to see if we can decide between selves.

4.2. Personal identity and economics

Before going to the philosophical discussion on personal identity, it is vital to analyse how economics perceive the topic of personal identity. In general, it is assumed that neoclassical economics deals with individual agents. Many economists think that they do not even need to discuss personal identity, because the axioms of neoclassical economics guarantee the existence of individuals (Davis, 1995). Two main features make it possible. Firstly, neoclassical economics is based on methodological individualism, which means that the whole analysis starts with individual agents. Secondly, neoclassical economists assume that people do not change but circumstances. The static character of the analysis enables economists to easily re-identify individuals over time. It is possible, because a static analysis investigates how people make choices in response to changes in incomes and prices (Davis,

1995). Moreover, economic agents are represented in terms of their preferences. This assumption seems to support the claim that economists' theory does not need to be involved in a discussion on personal identity, because preferences always belong to a given individual. Having a preference always presupposes some individual who does the preferring in question (Martin, 1987). This perspective was used by Becker when he claimed that preferences are stable and unchanging. For Davis, this assumption is the basis for individual identity in neoclassical theory, because it differentiates people from each other and makes them consistent over time (Davis, 1995).

It goes without saying that individuals' preferences cannot be willy-nilly. As it was indicated in the first chapter, they must be transitive and complete to be called rational. Thanks to these assumptions, it is possible to construct a complete, transitive, and continuous revealed-preference ordering from choices. The consistency of choice means that a rational agent (individual) will never make choices that undermine her own self-interest. In this way, economists can define self in terms of its own interests which can be represented by a utility function which is derived from well-formed preferences (Davis, 1995).

Of course, there is a lot of problems with rational choice theory, and we know that it is often the case that the assumptions of transitivity and completeness are violated. Behavioural economics is probably best known for indicating problems with these assumptions. Some of them were presented in Chapter 1 (menu-dependence, money pump, prospect theory). Therefore, in this chapter, the focus is on the economic assumptions about self. Mostly, on the very assumption that an economic agent has one self which is stable over time. To answer the question whether we can isolate self which is the same over time, we need to look at one of the most important debates in philosophy. The discussion on personal identity. Of course, the debate is extremely sophisticated and because of the book's scope and size limits, only some central concepts of this discussion are analysed.

To begin with, it is important to understand that there is no single problem of personal identity but a wide range of question connected with this issue (Olson, 2017).

Who am I? In general, personal identity refers to particular properties to which a person feels ownership or attachment. In this sense personal identity consists of those features that somebody takes to define her/him as a person.

Personhood. Answer the question what is it to be a person? What is necessary, and what suffices, for somebody to count as a person, as opposed to a nonperson?

Persistence. What does it take for a person to persist from one time to another to continue existence than cease to exist?

The main problem in the context of this chapter is personal identity over time.¹ It is especially important, because in neoclassical economics it is assumed that people have one self which is consistent over time. Thanks to this assumption, people can assess and compare their present and future preferences (the discounted utility model (Samuelson, 1937)). When we talk about personal identity over time it is necessary to address one common confusion. The persistence is a necessary but not sufficient condition for personal identity over time. So, sometimes because of the accident, Alzheimer disease, or a dramatic change in character we would be no longer the same person we were before. However, we would still exist (persist). Personal identity answers question like ‘What makes it true that a person at one time is the same thing as a person at another time?’ Personal identity is the unique numerical identity of a person in the course of time, viz., the necessary and sufficient conditions under which a person at one time and a person at another time can be said to be the same person, persisting through time (Korfmacher, 2018). The most illustrative example of the question about personal identity over time is the situation when one points to a child in an old class photograph and say ‘that’s me’. To claim that we need to answer numerous questions such as:

What makes you that one, rather than one of the others? What is it about the way she relates then to you as you are now that makes her you? For that matter, what makes it the case that anyone at all who existed back then is you? (Olson, 2017)

The simplest theory that explains personal persistence over time is continuous bodily existence. It assumes that the person is the same if it has the same physical body. However, the thought experiment known as the ‘ship of Theseus’ shows the problems in determining whether one physical body at one time is the same thing as a physical body at another time. Theseus took a long journey and with time parts of the ship worn out and they had been replaced. The question is how many parts of the ship must be replaced to conclude that the ship is no longer the ‘Theseus ship.’ It is a paradox, because one cannot point the exact moment in which it happens. Some can argue that even if all the parts of the original ship were replaced, it would still be a Theseus ship. Hobbes (1655, ch. 11.7) introduces the problem with this kind of reasoning. Imagine that someone gathers up all the original worn out parts of Theseus ship, restores them and builds the exact ship like the Theseus ship. Hobbes asked which ship, if either, would be the original ship of Theseus.

As has been pointed out, the concept of bodily continuity is very problematic in the case of material things. It is even worse when it comes to people. Over time

¹ It is a highly debated topic in philosophy, for more see (Kolak, & Martin, 1991; Noonan, 2004; Olson, 2017). Classic anthologies include Perry (1975), Rorty (1976), and Martin and Barresi (2003).

our bodies age and grow, gaining and losing matter. Each touch destroys part of our skin. In seven years all our cells are replaced by the new ones. Moreover, we can change our hairstyle, grow a beard or go to the gym to build some muscles. Some can argue, as in the case of Theseus ship, that we are still the same person, because we have the characteristic body features which sustain some continuity (you do not have a new body after one day in the gym). However, it is very problematic to base the persistence of personal identity on bodily features, because they are vague. As in the case of Theseus ship, we cannot exactly indicate a point where the SHIP is no longer the Theseus ship. This problem is called 'continuum fallacy' and is connected with 'Sorites paradox.' Imagine a bold man. If he had one piece of hair he would still be bald, the same goes for a second and third piece of hair. If he had 1000 hairs, he would be no longer bold. However, the problem is that we do not know the exact point in which he ceased to be bold (999 hairs means he is bold?). Because of these problems, it seems very problematic to ground the persistence of personal identity over time in the continuous existence of our bodies.

The second view on personal identity, which was very popular, is mental substance. This view is associated with Plato, Descartes, and also with many religious beliefs. In this approach, persons are immaterial souls. Thanks to this concept, people can survive their bodily death. Secondly, because our souls are immaterial and independent of the body, personal identity over time can be grounded on the persistence of non-physical substance, in spite of the continuous change of the body. Although this view is prevalent among religious people, it is strongly criticized by many philosophers of mind. In general, the mind-body distinction which was created by Descartes is perceived to be wrong and counterproductive (Dennett, 1969, 1991). The main objection towards mind-body theory is that it cannot explain how a non-material mind can influence the material body and *vice versa*. So, how the nonmaterial desire for food can cause a person to move and act to obtain this food? The biggest problem with an immortal soul is not lack of the connection between mind and body, but its undetectability. Even if we assume that immortal soul exists, we will have a problem to say anything about it. Therefore, we do not have any possibility to know, e.g., whether God destroys and replaces our souls every five minutes. Because of the undetectability of immortal soul, any discussion about personal identity is impossible.

The third main conception of personal identity was presented by John Locke and it is the most intuitive one. Locke (1689) considers personal identity (self) to be founded on consciousness (memory). We are the same person to the extent that we are conscious of the past and future thoughts and actions in the same way as we are conscious of present thoughts and actions. If consciousness is this 'thought' which "goes along with the substance [...] which makes the same

person," then personal identity is founded on the repeated act of consciousness: "This may show us wherein personal identity consists: not in the identity of substance, but [...] in the identity of consciousness" (Locke, 1689, pp. 242, 243).

Locke criticizes both mental and bodily concepts of personal identity. In the first case, somebody can say to be the reincarnation of Plato (having the same soul). However, for Locke one would be the same person as Plato exclusively if one had the same consciousness of Plato's thoughts and actions. Locke was against bodily criterion of personal identity, because the body can change, but the person remains the same. Locke's conception based on consciousness is also called psychological continuity and it is very intuitive for people. In the case of personal identity, philosophers very often use thought experiments to see what people think. Locke was not an exception and he proposed one of the first mind swamps thought experiment (Locke, 1689, p. 241). Imagine that the mind of the Prince enters the body of the sleeping Cobbler (whose own mind departs). The individual who later wakes up has the same body as the Cobbler but is a different Person (the Prince). In general, people think that if their brain had been transplanted into a new body, they would have got a new body rather than somebody else would have got a new brain. The intuition is that psychological continuity is the criterion for personal identity. Most people in Western countries feel drawn to psychological-continuity views (Nichols and Bruno (2010) give experimental evidence for this). Moreover, many philosophers are the supporters of psychological-continuity approaches (Garrett, 1998; Hudson, 2001; Johnston, 1987; Lewis, 1976; Nagel, 1986; Noonan, 2004; Nozick, 1981; Parfit, 1971, 1984; Perry, 1972; Shoemaker, 1970, 1984, 1997, 1999; Unger, 1990, 2000).

Although memory criterion is popular and intuitive, it is widely criticized. The first problem is connected with discontinuities in consciousness. Thomas Reid (b. 1710, d. 1796) presented the example of the Old General who remembers nothing of his boyhood (Reid, 1850, p. 235, 236). For Locke, this old general is not the same person as the boy. It does not make sense, because we know that they are the same person. For Locke, the old general is not the same as the boy since there are no direct psychological connections between them. It is true that the young boy has a connection with the young officer and that officer is connected with the old general. However, the old general has no memory of being a young child. Because this conclusion violates the transitivity of identity (young boy \neq old general), personal identity relations cannot consist in direct memory connections. However, it is possible to resolve this problem. Instead of direct connection, we can refer to indirect chains of psychological connections. In this case, the old general is identical with the young boy, because the old general remembers being the officer and the officer remembers being the young boy.

The next problem connected with discontinuities in consciousness is that we lack memory of periods of sleep or that we do not remember what we ate for

breakfast a few days ago. It would be insane if it did mean that we ceased to exist because of this. To solve this problem, one needs to appeal not only to memory but to psychological aspects as well. Therefore, we do not remember what we ate a few days ago at breakfast, but we remember what kind of person we were there.

The next considerable problem with Locke's theory is that we cannot know if our memories are genuine rather than false.² To assume that these memories are ours, we need to presuppose that there is a YOU. It was Butler who spotted that Locke made a "wonderful mistake," because he failed to recognize that consciousness presupposes identity, and thus cannot constitute it (Butler, 1975, p. 100). I can remember only my own experiences, but it is not my memory of an experience that makes it mine; rather, I remember it only because it is already mine. Therefore, while memory can reveal my identity with some past experiences, it does not necessarily make that experiencer me. What I am remembering, argues Butler, are the experiences of a substance. The same substance that constitutes me now (Shoemaker, 2016).

Probably the biggest problem with psychological continuity approaches is visible in the cases where we are presented with the possibility of two (or more) future persons that are psychologically continuous with a presently existing person. Can one really become two? The most famous example of this problem was investigated by Parfit on the example of teleportation (Parfit, 1984, pp. 178–190). In sci-fi movies functioning of 'teletransporter' is simple. The person copy is created atom by atom and then teleport to another place with 'teletransporter' while the original person is destroyed. Some argue that teleportation is a way of transportation, whereas others argue that we die in the process. Parfit gives the example of teleportation from Earth to Mars. Imagine that you go to 'teletransporter,' push the button and nothing happens. The service staff tells you that some technical problem occurred. Yours exact copy is on Mars, but your Earth version has not been destroyed. However, you should not worry, because you on Earth will be destroyed in five minutes and your copy on Mars will exist.

The example of transportation demonstrates that both bodily and psychological criteria are not enough for personal identity. For five minutes we have two same bodies with the same memories and personalities. You on Earth are the same person as you on Mars. However, it is impossible for you two to be the same person (have the same identity) because of the transitivity criterion which states that for two things to be identical they must have the same properties. In the case of transportation, this criterion is not fulfilled, because two copies are in

² It is not as hypothetical example. Often people remember some things that did not happen. This psychological phenomenon is called 'false memory' (Brainerd, & Reyna, 2005).

different places in space. This example shows that personal continuity is not a sufficient criterion for personal identity. Moreover, in the case of teleportation there is no right or wrong answer to the question whether somebody has ceased to exist. Which of these two copies is me? Me on Earth or me on Mars? If my copy on Earth dies, will I exist as the same person on Mars?

Psychological continuity is not sufficient for the criterion of personal identity, because one can at least imagine the situation in which two or more people have the same psychological features and memories. Despite this fact, Parfit thinks that the existence of a person consists in nothing over and above the existence of a brain and body and the occurrence of the interrelated series of mental and physical events. However, persons themselves are distinct from their bodies and psychologies. To understand this seemingly inconsistency we can use the analogy. Cellini's Venus is made of bronze. Although the lump of bronze and the statue itself exist, these objects have different persistence conditions. If melted down, Venus ceases to exist, while the lump of bronze does not. On the other hand, a lump of bronze alone is not Venus. Therefore, they are not identical. The lump of bronze constitutes the statue. The same is true of persons, who are constituted by, but not identical with body and brain (Korfmacher, 2018).

What Parfit takes from the teleportation paradox is that we should dismiss the concept of personal identity altogether. For Parfit personal identity is all or nothing approach in which there is no place for the intermediate cases like teleportation in which the concept of personal identity does not work (the problem of transitivity). In conclusion, what matters is the survival of the person, which is connected with psychological continuity and not absolute persistence of personal identity. Therefore, in the end, Parfit dismisses the traditional concept of self and he argues for the conception of survival, which is connected with psychological continuity in which different degrees of personhood are possible.

The fourth main conception of personal identity is the illusion of self. Hume was the influential supporter of this approach.³ He attacks the idea that there is some spiritual substance which each person can detect in oneself. Hume argues that he was unable to detect this substance in himself. Because there is no spiritual substance, Hume concludes that "man is a bundle or collection of different perceptions which succeed one another with an inconceivable rapidity and are in perpetual flux and movement" (Hume, 1896, p. 174), where perception is a way of designating items of consciousness (Davis, 1995, p. 43). Hume thinks that we do not have self, because we change all the time. Every experience makes us different. Despite the constant changes in themselves, people think

³ It is necessary to notice that more than 2000 years before Hume, Buddhism indicated the same thing. Buddha argues that self is an illusion, because we are built on many different and constantly changing elements (see (Siderits, 2019)).

that they are the same person they were three years ago. People understand that their features change, but they think that they change without changing the self. For Hume, however, there is no distinction between the different features of a person and the self. If features change, the self changes as well. This is why he argues that people do not have a distinct and unique self.

Hume understood that people mistakenly assumed that they “possessed of an invariable and uninterrupted existence through the whole course of our lives” (Hume, 1896, p. 174). One of his goals was to explain this mistake. For him people misunderstand personal identity, because they confuse two ideas that must be treated separately. One is an object that endures change and the second is a succession of closely related objects. Hume argues that people are mentally lazy and they conflate the latter idea with the former. They perceive a collection of closely related objects as one thing. This generates a habit of thinking where people overlook small differences and gradual changes in things. Thus, people see a connection between themselves now and three years ago, because they do not see thousands of various small changes, but one pattern of change.

Hume accused the other philosophers of starting speculation about personal identity. He argued that they knew that people mistakenly perceived themselves as one self, but instead of destroying this illusion they invented a metaphysical ground for self (Davis, 1995). Hume’s criticism of self, however, has problems of its own. He argues that he could not detect the spiritual substance in himself, and because of that we are only bundles of perception. However, one can ask, how did Hume know that he was looking into himself and not that he was looking into the minds of others when surveying his perceptions? Hume’s approach seems to presuppose that one possesses some means of distinguishing oneself for another person. Moreover, to see that the different processes and states of mind are unfitted, there has to be something which perceives their unity (a perceiver). The existence of this would be no less mysterious than a personal identity (Davis, 1995).

Nowadays, many neuroscientists follow the steps of Hume and Buddha and argue that self is an illusion (e.g., (Harris 2012; Wegner, 2002)). Much of the research is devoted to non-consciousness. For example, Wilson (2002, p. 24) observed that the brain of an individual receives through the five senses more than 11,000,000 pieces of information every second, whereas it is capable of processing consciously about 40 pieces of information. What happens to the other 10,999,960 is the most crucial question here. This research poses a threat to our perception of humans who are governed by rational consciousness. Over the years it was the feature that has differentiated people from animals with their visceral motives. Nowadays, both psychology and neuroscience demonstrate that also for humans ‘most of moment-to-moment psychological life must occur through nonconscious means’ (Bargh, & Chartrand, 1999, p. 462).

Bargh and Chartrand (1999, p. 462) identify “an automatic effect of perception on action, automatic goal pursuit, and a continual automatic evaluation of one’s experience”. Although people are characterized by consciousness and rationality, they acquire information, choose, and evaluate it in non-conscious ways (Zajonc, 1980, 2000).

In the end, it is necessary to underline shortly how huge influence society has on constituting individual selves. In value-free economics, society does not play a significant role. People have their own preferences and even if they take other people into consideration they do it only because it impacts their own utility. Baumeister (1998) goes deeper into the analysis of the individual by considering three roots of selfhood: the agentic self, the reflexive self, and the social self. Especially the last one is important. Nowadays, in the individualistic culture, we tend to forget how vital society was for people. The philosopher Boethius (b. 480, d. 524) was the first to trace the genealogy of the concept of the person. The word *persona* was originally used in Latin to designate the mask worn by actors in the theatre. Moreover, Romans used the term ‘person’ for every person who they recognized from her or his face (Ballet, Bazin, Dubois, & Mahieu, 2013, p. 44). Over the last years, Sen was the economist who argued that what people do may depend on their social identity which is constituted by their membership in different social groups. Sen (1990, pp. 125, 126) points to data suggesting that Indian women perceive themselves more as members of their families rather than as individual selves and it has an influence on their choices.⁴ People may also identify with the firm where they work, different associations and clubs, nation, caste, religion, and so forth. We tend to forget that the concept of individualism is something relatively new in our history and even now it is highly praised mainly in the Western World. Goffman (1959) presents comprehensively that the birth of individualism started with a Renaissance in the 14th century. Before that, people had identified themselves with tribal group and family.⁵ It must be underlined that other people still may enormously influence our identity and behaviour (Fiedor, & Ostapiuk, 2017). Barkley’s sentence *esse est percipi* can be used to argue that to be a person, we need to be perceived by other people and to conclude that “no man is an island” (Donne, 1624). We are social animals (Aristotle) and in general our self-awareness comes from the fact that we know that we can be observed by other people, which was also underlined by Smith (1759). Of course, economics many years ago demonstrated how important society was (institutional economics). Unfortunately, neoclassical economics did

⁴ In these situations, social values can be perceived as metapreferences.

⁵ Elias (1991, p. 101) draws a connection between changes in books and people’s perception of individuality. Until the second half of the 19th century writers mainly described what people did or what happened in general. Later the focus has shifted on what people experience, which the extreme case is the ‘stream of consciousness.’

not seem to take these insights into serious consideration. Thankfully, over the last years, the awareness of social influence on people has risen in economics mainly thanks to new institutional economics (Godłów-Legiędź, 2003, 2010; Hodgson, 1998, 2000, 2007; Ratajczak, 2005, 2011; Wilkin, 2016).

The discussion on personal identity and the problems connected with the perception of one self over time obviously does not exhaust the topic in any way. The discussion was presented to highlight a few points. Firstly, the discussion on self is significant in the context of people's responsibility for their actions. In neoclassical economics, it is assumed that people have discerned self and free will which makes them responsible for the choices they make. (This assumption will be especially important from the perspective of the last chapter of this book.) The next goal was to show that neoclassical economics' assumption that a person has one self with stable preferences is false and one's self is changing all the time. This goal is crucial, because it can be argued that the concept of multiple self, which is examined in this book, should not be treated seriously, because it is only a metaphor. The discussion on personal identity should demonstrate that the concept of multiple self is much closer to reality than the concept of one, unchanging self that is used in neoclassical economics. This discussion should serve as a justification for further discourse on different preferences that people can have. The last goal of the analysis of personal identity was to indicate that preferences do not establish a conception of individual as a distinct being as it is assumed in neoclassical economics. Hume's argument was used to show the problems with this conception. It was presented that we can never detect anything resembling the self in our own preferences (Davis, 1995). The only thing that we can detect are the contents of consciousness or preferences. One's self is not a content of consciousness as well as preferences. To see that point more clearly we can get back to Locke's argument for the connection between memory and self. In the case of economics, we can also observe an inherent circularity when we assume that preferences pick out their owner. Davis concludes that "rather than constitute a basis for personal identity, claiming preferences are own preferences seems to presuppose it" (Davis, 1995, p. 43).

As Davis (1995, 2003) argues, the axiomatization of preferences creates a sophisticated but artificial construction which only organizes preferences to allow economists to pick out and identify individuals. In the end, however, this axiomatization has to presuppose rather than demonstrate that there are enduring and unique individuals to whom well-formed preferences are attached. Neoclassical economics lacks a coherent theory of the individual because it fails to do this.

4.3. The conception of multiple self

*Two souls, alas, are dwelling in my breast,
And one is striving to forsake its brother*

(Goethe, 1961, p. 145)

In the second chapter, the conception of multiple self was shortly presented in the context of the discounted utility model (Samuelson, 1937). Because of the hyperbolic discounting, we know that people's preferences are not always time consistent and that people are too much focused on the present (present bias). This idea undermines revealed preference and rational choice theory. In this chapter, the concept of multiple self is developed. It is presented that there are many multiple self conceptions with different assumptions. Moreover, the idea of precommitment (presented briefly in Chapter 2) is used to justify the phenomenon of weakness of will, which undermines revealed preference theory. However, the most crucial task of this chapter is to demonstrate that comparing decisions over time by utility measurements is very problematic. Therefore, economists will need to decide normatively which values are more important.

Firstly, it is essential to underline the fact that over the last years the conception of hyperbolic discounting has been gaining influence predominately in psychology but also in economics and philosophy. However, the problems with intertemporal choices are not something new. Plato quoted Socrates to present what can go wrong when people assess the future:

Do not the same magnitudes appear larger to your sight when near, and smaller when at a distance? [...] Is not [the power of appearance] that deceiving art which makes us wander up and down and take the things at one time of which we repent at another? [...] Men are in their choice of pleasures and pains, that is, in their choice of good and evil, from defect of [...] that particular knowledge that is called measuring. (as cited in Ainslie, 2001, p. 4)

Aristotle called this behaviour *akrasia* which is also known as weakness of will.⁶ Also, Freud (1956a, 1956b) underlined the differences in preferences which depend on time. He made a distinction on processes that serve the long-range goals (reality principle) and those which serve the short-range goals (pleasure principle).

What is important here is that having different, even contradictory, preferences is not an anomaly. Frequently people do things that they did not want to do

⁶ For now, both terms are used interchangeably even though some philosophers make distinction between these two terms (e.g., (Holton 1999)). The reason for this simplification is clarity for a reader who does not need to be entangled in philosophical discussion at this point.

(Jeffrey, 1974). Often, we hear after such choice “I know it’s irrational, but...” It poses a problem for neoclassical economics which assumes that people have consistent and well-defined preferences and that they choose what they want to choose. However, it looks more like individuals often have contradictory preferences. This violates the essential, from neoclassical economics perspective, assumptions of transitivity and completeness of preferences (WARP). However, in real life it is not always a problem. Sen (1974, pp. 378, 379) writes that “there is nothing particularly schizophrenic in saying: ‘I wish I had a vegetarian’s tastes, for I disapprove of the killing of animals, but I find vegetarian food so revolting that I can’t bear to eat it, so I eat meat.’”

These differences in human preferences and motivations have been fascinating intellectuals for ages. There are many explanations and concepts concerning multiple selves. However, not many of them have been particularly important from the perspective of economics. This is the reason why only a few conceptions are presented here. Moreover, many of multiple self conceptions have not been formalized. Elster (1987) presents a summary of multiple selves conceptions.

The loosely integrated self. Elster (1987, p. 3) suggests that multiple selves often ‘turn out to be little more than failures of coordination and integration,’ such as an individual who has motivations and beliefs that contradict each other. In this conception people can realize these contradictions and they try to revise their beliefs to be consistent with motivation.

Faustian selves. They refer to famous ‘two souls’ within Faust. One part of the person has higher order preferences, whereas the other part has desires that clash with them. The conflicts between Faustian selves can be cursory (can be resolved by deliberation) or deeply ingrained and last for years.

Hierarchical selves. Elster (1987, p. 11) argues that Faustian selves can be expanded. We can have more than two selves which also can be hierarchical. For example, metapreferences and second-order desires that were presented in Chapter 2.

The Freudian legacy. Freud’s theory presents a famous distinction on the ‘id’, ‘ego’, and ‘superego’ (Freud, 1956b). Elster (1987, p. 20) identifies these three concepts as ‘agents’ that have different tasks in the person. These agents work in a different framework (conscious, preconscious and unconscious). What is important about this concept is that it helps to understand some inner processes and conflicts that can exist in a person for years.

Split brain – split mind. This concept is connected with the finding that two hemispheres of the human brain can operate independently from each other. It gives rise to many thought experiments which were used in various theories of personal identity over time. Elster (1987, p. 23 ff.) argues that it is not the best

concept to describe multiple self theory. It is beyond this book's scope to develop this idea further because the nature of this discussion is purely biological. However, a compelling argument against split brain – split mind is a fact that different hemispheres are responsible for different cognitive tasks but they do not have different preferences.

Psychologists are the most interested party in the discussion about multiple self. Nowadays, it is commonly assumed that there are two completely different processes that influence human mind. The first type has been variously named “controlled” (Schneider, & Shiffrin, 1977), “reflective” (Lieberman, 2003; Strack, & Deutsch, 2004), “rational” (Epstein, 1994) or just “system 2” (Kahneman, 2011). This process is characterized as conscious, deliberative, effortful and can be identified with *homo economicus* (Thaler, & Sunstein, 2008). The second type of the decision process is called “automatic” (Schneider, & Shiffrin, 1977), “impulsive” (Strack, & Deutsch, 2004), “reflexive” (Lieberman, 2003), “experiential” (Epstein, 1994) or just “system 1” (Kahneman, 2011). The processes in this category are fast, unconscious, effortless, and automatized.

What differentiates psychologists models from some multiple self theories is that in the former, two different psychological processes work together. For instance, Rottenstreich and Hsee (2001), and Hsee and Rottenstreich (2004) state that preferences are both influenced by deliberative valuation by calculation and an affective valuation by feeling. Both economists (Loewenstein, & O'Donoghue, 2004) and psychologists (Metcalf, & Mischel, 1999) propose the models where people's behaviours are co-determined by an automatic, 'hot' affective system and an effortful, deliberative 'cold' cognitive system (see (Alós-Ferrer & Strack, 2014)). It is important to keep in mind that in spite of the fact that human brains have completely different cognitive processes, they are interconnected. Thanks to this fact we can agree with Loewenstein and his claim that we “do not believe that there are little selves in people with independent motives, cognitive systems, and so on” (Loewenstein, 1996, p. 288). When we discuss the multiple self concept we do not have in mind two *homunculi* that are in people's heads and have different personalities. However, it is literally possible to have more than one personality and this affliction is called 'dissociative personality disorder.' When the concept of multiple self is used in this book, it is used in a metaphorical sense. It is presented to signify that people have contradictory preferences that cannot be easily compared with each other within one unified utilitarian framework. In this chapter, the main focus is put on the conception of multiple self from the perspective of choices over time, and the fact that hyperbolic discounting indicates that there are different selves which make different choices depending on the time horizon. This concept of multiple self is fundamental in the context of neoclassical economics because it undermines the discounted utility model (Samuelson, 1937). Moreover, it shows that the revealed preference theory does

not work perfectly because people's choices do not always reveal their real preferences. Hyperbolic discounting undermines the rational choice theory because it demonstrates that individuals' preferences are time inconsistent. The discussion on the concept of multiple self will also be investigated in the context of weakness of will in which agents behave irrationally and act against their best interest, especially when choices are made over time.

4.4. The conception of multiple self over time

Economists were not much interested in the problem of time inconsistency until the 1980s, with the exception of Strotz (1955). It was psychologist Ainslie who introduced the idea of hyperbolic discounting (1975). Thaler and Shefrin (1981) implemented this idea in economics. It gained appreciation with time and economist had to admit that hyperbolic discounting is a common feature of human behaviour. However, economists dealt with inconsistent time preferences in a specific way. They created the models where choices over time are compared by utility measurements by taking into account hyperbolic discounting among other manifestations of bounded rationality like self-control, temptation, dual-self, etc. (Andreoni, & Sprenger, 2012; Brocas, & Cartillo, 2008; Cohen, Ericson, Laibson, & White, 2020; Dasgupta, 2008; Fudenberg, & Levine, 2006, 2012; Gul, & Pesendorfer, 2001, 2004; Hoch, & Loewenstein, 1991; Laibson, 1997; Loewenstein, & O'Donoghue, 2004; O'Donoghue, & Rabin, 1999, 2001). Survey of the intemporal models can be found in (Ericson, & Laibson, 2019). How these models work will not be elaborated separately, because they face the same problems as the static models (the discounted utility model (Samuelson, 1937)) which will be analysed below. Namely, it is impossible to compare decisions over time exclusively by means of utility measurements.

Schelling was one of the economists who noticed this problem. In Chapter 2, Schelling was mentioned in the context of precommitment strategies (the example of Odysseus) which are used by people because of the contradictory preferences over time. As it has been indicated before, people use these strategies of their own volition to remove future options from which they choose because they are afraid that they succumb to the temptations when the choice comes. For Schelling this behaviour was really disturbing when he referred it to neoclassical theory of rational choice and revealed preferences. He indicates that people frequently choose between the long-term goals and the short-term pleasures and sometimes they curb their freedom to achieve long-term goals. Schelling (1984, pp. 1, 2) enumerates examples of this behaviour:

Please do not give me a cigarette when I ask for it, or dessert, or a second drink. Do not give me my car keys. Do not lend me money. Do not lend me a gun. Besides denial there are interventions. Do not let me go back to sleep.

Interrupt me if I get in an argument. Push me out of the plane when it's my turn to parachute. Don't let me go home drunk unless you can remove my children to a safe place. Blow the fuse if you catch me watching television. Make me get up and do my back exercises every morning. Keep me moving if I am exhausted in the wilderness. Pump my stomach if you catch me overdosed with sleeping pills.

For Schelling the fact that people use precommitment strategies was a demonstration that multiple selves exist. Thanks to the multiple self conception, Schelling was able to explain precommitment strategies and their existence was not a problem for him as it was for value-free economics. What was important for him was a question which self should be in control. The self which has long-term preferences or that which falls into the temptation? To answer this question, Schelling (1984, p. 9) analyses the example of captain Ahab from *Moby Dick*, who lost his leg because of a bad wound:

The blacksmith enters with a hot iron to cauterize the stump. Ahab begs not to be burned. The crewmen hold him down as he spews out the apple in a scream, and steam rises where the iron is tormenting his leg. The movie resumes with Ahab out of pain and apparently glad to be alive.

At first, Schelling treats this example as an easy situation. We get a lot of utility in exchange for our limited freedom. Life in exchange for short but a very intensified moment of pain. Later, however, he notes that even if Ahab asked the smith to cauterize his wound and even if felt grateful for the operation which was performed upon him, it would not mean that this was the best decision which gave him the greatest utility. Schelling (1984, p. 9) writes: "If you burn me so that I may live I'll thank you, afterward, but that is because I'll be feeling no pain and not anticipating any when I thank you". He concludes by saying:

How do we know whether an hour of extreme pain is more than life is worth? Alternatively, how do we know whether an hour of extreme pain is more than death is worth? The conclusion that I reach is that I do not know, not for you and not for me. (Schelling 1984, p. 9)

Schelling stated that it is impossible to assess which choice is better if we only take utility into an account. He claims that every moment is different, and it is impossible to compare different periods of time. Thus, Schelling reached a similar conclusion to neoclassical economics and Becker. It is impossible to compare choices in terms of utility, because utility is absolutely relative. As far as neoclassical economics is concerned, it is thought that if we cannot assess whether some actions give more utility than other, we assume that people are rational and always do things that maximize their utility. If we did not assume that, the theory of revealed preference would not work. Consequently, we would be back to the situation when neoclassical economics had no reliable methodology.

Schelling comes to another conclusion. He thought that it is impossible to compare utility over time, because:

Each self is a set of values; and though the selves share most of those values, on the particular issues on which they differ fundamentally there doesn't seem to be any way to compare their utility increments and to determine which behavior maximizes their collective utility. (Schelling, 1984, p. 8)

We cannot use utility measurements because human beings have two different selves and they have different preferences. Thus, economics should handle these selves separately. Schelling did not want to decide which of these two selves is more important because we are not able to compare their utilities. However, this relativity is not a problem for Schelling, because:

Sometimes, but not always, it is easy to know which is Dr Jekyll and which is Mr Hyde. The person who drinks and becomes vicious, or a bore, and is morose about it for days afterwards; the person who continually resolves to demand that increase in pay and never musters the courage; and the person who walks into a casino for a little sociable gambling, loses more than he intended, commits more to recover it, and emerges traumatized after blowing his bankroll, all seem to present an unequal pair, a 'straight' ego and a wayward alter. (Schelling, 1980, p. 98)

However, Schelling was aware that he did not have a framework to decide which self is more important. He wrote that "anyone who is happily addicted to nicotine, benzedrine, valium, chocolate, heroin, or horse racing, and anyone unhappily addicted who would not elect the pains and deprivations of withdrawal, are not my subject" (Schelling, 1984, p. 4). However, he made a subtle distinction which helped him to justify a choice:

I am not concerned with whether cigarettes or rich desserts are bad for you, only with the fact that there are people who wish so badly to avoid them that, if they could, they would put those commodities beyond their own reach. (Schelling, 1984, pp. 4, 5)

Even though he did not want to choose, the last quotation shows that in the end, he chooses reason and the 'long-term human' over a man focused at present moment. His slight emphasis on people who succumb to temptations against their better judgment is known as *akrasia* or weakness of will. Economists do not recognize *akrasia*, because it is assumed that people do what is best for them (in accordance with revealed preference theory). Thaler presents why neoclassical economics resists that the phenomenon of *akrasia* exists.

In the 1970s he organized the supper for his co-workers from Chicago who were known for their orthodox approach and who assumed that people are fully rational. Thaler was a good host and he prepared a bowl of cashews for his guests to whet their appetite before the supper. In a short time, it occurred that guests too often reached for the snack and Thaler took away a bowl of

cashews, because he did not want the guests to overeat before the supper. His decision was met with guests' approval. They wanted to eat only a few cashews, but they gave in to the temptation and ate more than they thought was prudent. The guests' approval does not make sense from the perspective of neoclassical economics because people are rational agents and know what is best for them. Therefore, having more choices is better, because individuals can find a more preferable option. Thaler (2015, pp. 94, 95) presents a hypothetical discussion between rational *homo economicus* (econ) with a human being about this very topic:

ECON: Why did you remove the cashews?

HUMAN: Because I did not want to eat any more of them.

ECON: If you did not want to eat any more nuts, then why go to the trouble of removing them? You could have simply acted on your preferences and stopped eating.

HUMAN: I removed the bowl because if the nuts were still available, I would have eaten more.

ECON: In that case, you prefer to eat more cashews, so removing them was stupid.

This discussion leads nowhere. Neoclassical economics cannot admit that weakness of will exists, because it would imply the dismissal of revealed preference theory.

4.5. Weakness of will

To answer the question whether weakness of will (*akrasia*) exists, we need to investigate the philosophical discussion. In general, *akrasia* is perceived as doing something against one's better judgment. So, we do something and we know it is not the best course of actions for us to take. Not only were the economists perplexed about this possibility, but from the very beginning, philosophers have discussed this problem as well. In Plato's *Protagoras*, Socrates denied the possibility of weakness of will: "No one," he declared, "who either knows or believes that there is another possible course of action, better than the one he is following, will ever continue on his present course" (Plato, 1961, p. 348). In the 20th century Hare agreed with Socrates. He argues that if one evaluates something to be the best action for one, then one must do it. This is a case because evaluative judgments are connected with a prescription (Hare, 1952, 1963). Therefore, when one judges that the best action for one is to quit smoking then one must quit smoking.

Often people say that they know that quitting smoking would be the best action for them to take but they cannot do it. For Hare (1952, 1963) as well as for Socrates it does not mean that weakness of will exists. In this case, one cannot quit smoking, because deep down one still wants to smoke and this is one's

strongest desire, not desire to quit smoking. We can see how similar this line of reasoning is to value-free economics. In Chapter 2, Becker's approach towards this issue was presented:

The claims of some heavy drinkers and smokers that they want to but cannot end their addictions seem to us no different from the claims of single persons that they want to but are unable to marry or from the claims of disorganized persons that they want to become better organized. What these claims mean is that a person will make certain changes, for example, marry or stop smoking – when he finds a way to raise long-term benefits sufficiently above the short-term costs of adjustment. (Becker, & Murphy 1988, p. 693)

As has been indicated before, for Becker declarations do not mean anything. They are only a cheap talk. If you really want to quit smoking, you will do it. Thanks to this line of reasoning, present bias is not a problem in the economic approach. It is not an irrational feature of human behaviour that indicates that people do not do what they really want to do. They are just overpowered by present desires and deep down they want them. The tautological (axiomatic) construction of the economic approach in which rationality is perceived instrumentally enables economists to deny the possibility of weakness of will and still be consistent. Despite the approaches in philosophy and economics, where the weakness of will is impossible by definition, for many people it is a real feeling. If people always do what they want deep down how is it possible that so many of them feel regret after choice? The contemporary debate about this phenomenon started anew thanks to Davidson (1980).

Davidson thinks that weakness of will is possible and he is mostly known for providing the definition of akratic action which is commonly used. In doing x an agent acts akratic if and only if: (1) the agent does x intentionally, (2) the agent believes there is an alternative action y open to him, and this guarantees that the action is free, (3) the agent judges that, all things considered, it would be better to do y than to do x (Davidson, 1980, p. 22). It is important to underline the phrase 'all things considered.' It does not mean that if one thinks that an apple is more healthy than chips, this is one's evaluation. Here one only sticks to one set of values (health) versus another (taste). 'All things considered' means that one compares a different set of values and then makes a decision.

Nowadays, most philosophers interested in the weakness of will agree with Davidson that akratic action is not only possible but actual, and it is a common feature of people's lives. Moreover, most philosophers agree that *akrasia*, if possible, must be irrational (Stroud, & Tappolet 2003, p. 5). Here it is important to underline the distinction between economics and philosophy in the case of rationality. As was indicated before, neoclassical economics treats rationality in an instrumental way, whereas many philosophers tend to treat it as something which is worth investigating and striving for. Therefore, most neoclassical

economists do not assess whether some desires are rational or not, whereas philosophers ask these questions. Moreover, neoclassical economists analyse only human actions and they do not ask questions about people's motivations, whereas philosophers use a distinction between belief and choice in which belief is beyond observation (revealed choice). This different perception of rationality results in the situation where philosophers attempt to define what the rational behaviour might consist in, and economists do not.⁷

Thanks to this short introduction we can analyse whether *akrasia* is rational or not. As was stated before, *akrasia* refers to behaviour in which one chooses to do something that one believes not to be the best option for one. The belief that a certain option is not a good thing to do is either a rational belief for one to hold or not. When it is not rational belief to hold, then one is irrational in one's belief. For example, Martha is dissatisfied with her vacation in Spain and she believes that she should have gone to Italy instead. However, she is mistaken and the vacation in Italy would be much worse. Then, on the one hand, her belief in Italy as a better place for a vacation is false and it is irrational to hold it. On the other hand, if her belief that Italy is a better place for a vacation than Spain is true, then she has a rational belief to hold. However, in this case, it cannot be rational for her to go to Spain. Therefore, *akrasia* necessarily involves irrationality, either in her beliefs or in her choices (Wedgwood, 2003, p. 216).

Although many philosophers agree that *akrasia* must be irrational,⁸ they argue whether it is possible intentionally and freely to act against one's better judgment. For example, Watson (1977) argues that akratic action does not exist because acting against one's better judgment is not free. It is so because there is no clear difference between *akrasia* and compulsion in situations when the agent is motivated by emotion or desires that one is unable to resist. Watson considers two possible explanations of why people fail to resist a desire that they do not want to have and he rejects both of them. People's failure cannot be explained in terms of what they choose to do, because a choice must follow better judgment (Hare, 1952, 1963). It cannot be also explained by an insufficient attempt to resist. If action is free, we must assume that an agent is able to control oneself and here the question why one did not try hard enough arises. Watson concludes that the agent was unable to resist the rebellious desire and in both cases one's action was unfree. However, many philosophers argue that it is possible to distinguish between weak-willed and compelled action and Watson's conclusion is controversial.⁹

⁷ Some economists define rationality (e.g., ecological rationality by V. Smith, procedural rationality by H. Simon). However, they are treated as heterodox (Berg, & Gigerenzer, 2010) and instrumental rationality dominates in economics.

⁸ However, not everyone agrees that *akrasia* must be irrational, see (Audi, 1990; McIntyre, 1990).

⁹ For more discussion see (Bermúdez, 2018; Kennett, 2001; Mele, 1987, 2010; Tenenbaum, 1999; Wallace, 1999).

After setting the stage for weakness of will, we move to the most important goal of this chapter – providing a framework to decide between selves. Holton (1999) argues for the position similar to Schelling’s precommitment. For Holton weakness of will means that certain kind of intention – a resolution – was violated. The term ‘resolution’ refers to a plan or intention one has to hold in the face of potential temptations. So, we can identify resolution with precommitment strategies. Holton (1999, p. 52) thinks that “central to the idea of weakness of will is an overreadiness to abandon one’s resolutions”. Therefore, we can know that people are weak-willed when they try to fight with some unwanted desires by sticking to resolutions. For example, Janet needs to study for the forthcoming exam. She dismissed the proposal of her friends to go to the pub in the evening, because she wants to study (resolution). However, after 30 minutes of studying, she was tired and watched a movie instead. In the end, she failed to study but at least she tried to do so and she was able to dismiss her friends’ invitation. For Holton, this kind of behaviour indicates weakness of will.

Moreover, he makes a distinction between weakness of will (overriding one’s resolution) and akratic action (overriding one’s judgment). This distinction was criticized by some philosophers (see (Mele, 2010)) but it is a very useful distinction for the purpose of this chapter. Holton (2003, p. 52) writes:

Imagine someone who is convinced that all the arguments point the same way: he should give up meat. Yet suppose he does not. He is, therefore, akratic. But is he weak-willed? That, I contend, depends on other factors. Has he repeatedly vowed to give up, only to find himself succumbing time and again in the face of rare steaks and slow-cooked offal? Or does he unblushingly affirm while conceding practical inconsistency – that he has never had any intention of giving up, and never will? In the first case we surely would accuse him of weakness of will; in the second I think we would not.

This distinction is the same distinction which was made by Schelling when he argued that we should help only people who act against desires they do not want to have (by precommitment). For Holton and Schelling it is not enough to say that one does not want to smoke, one needs try to quit smoking. Otherwise, it can be only a cheap talk. The distinction between *akrasia* and weakness of will is useful because in the case of *akrasia* we do not know whether somebody really wants to quit smoking or one only says so. When one acts against unwanted desires and form a resolution, it is signal that this person really wants something else. It also provides a better framework than a vague framework used by many behavioural economists who rely on ‘New Year’s resolution test’ (Rizzo, 2016; Sugden, 2018).

Holton uses the idea of resolution to explain why it is impossible for some philosophers to accept the existence of weakness of will. They cannot accept it because they use the ‘Humean account’ (belief-desire account) where all

intentional actions are explained in terms of an agent's beliefs and desires. The agent simply acts on the desire that is the strongest. In this approach, resolutions are reducible to beliefs and desires (Holton, 2003, p. 51). Therefore, when someone tries to quit smoking but fails to do so, it means that deep down one still wants to smoke. Intentions have the same value as desires because in the end "reason is, and ought only to be the slave of the passions" (Hume, 1896, p. 213). This reasoning is very close to neoclassical economics because of instrumental rationality. Holton is a supporter of The 'Augmented Humean account' (belief-desire-intention account) in which beliefs and desires are not enough. He writes that "intentions, of which resolutions are a species, should be seen as a third mental kind, irreducible to the other two" (Holton, 2003, p. 51). This approach works similarly to the Humean account in respect of the end result. Therefore, "if a resolution is stronger than any contrary desires, the agent will stick to it; if the contrary desires are stronger, then the agent will act on them instead" (Holton, 2003, p. 51). However, the qualitative difference between desires and intentions makes weakness of will possible and it also explains why people feel regret after following unwanted desires. In the context of economics, intentions can be perceived as people's real preferences.¹⁰

The qualitative distinction between desires and intentions is necessary in the context of intertemporal choices when we need to choose normatively which self should decide. It cannot be done by utilitarian measurement which is used by contemporary economists where desires and values are compared within the same framework as in the Humean account. The limitations of the utilitarian framework are visible on the example of Mill who tried to answer criticism of utilitarianism which in his times was called 'pig philosophy.'¹¹ Mill differentiated higher pleasures from the lower ones with regard to quality, not using only quantity or intensity. He argued that "some kinds of pleasure are more desirable and more valuable than others" (Mill, 1863, p. 11). However, very often, we choose lower pleasures. Sometimes we prefer lying on the couch and watching comedy to reading Shakespeare. We crave for undemanding activities even if we do not place great importance on them. At first sight, it looks contradictory to Mill's theory when people want higher pleasures which are connected with intellect, not with the trivial thrills. However, Mill answers this criticism and explains the difference between human choosing in the short-term and the long-term perspective. It is true that we prefer lying on the couch and watching a comedy to reading Shakespeare, because this less demanding activity can give us more pleasure. However, people value Shakespeare more and it is not only

¹⁰ However, the idea that we can analyse people's preferences, which are beyond the revealed choices, is very controversial and is the matter under discussion in economics (Caplin, & Schotter, 2008).

¹¹ This term was famously used by Carlyle (1850, p. 260).

because they are taught to have respect to the classic. As far as Mill is concerned, the true choice is in the long-term perspective when individuals' preferences are changed and people look for things which are more valuable than simple pleasures. Despite wanting to lie on the couch and watch a comedy, if somebody asks people what they want to do in their whole life, the majority will choose Shakespeare or the other kind of high culture. Mill explains that we choose the lower pleasures even if we want the higher ones, because "occasionally, under the influence of temptation, postpone them to the lower" (Mill, 1863, p. 14). In other words, it means that we know the difference between comedy and Shakespeare but sometimes we succumb to our weaknesses. We find *Hamlet* a great play not because we like it more than the other pleasures, but because it challenges us and makes us better people.

In this context, Mill unknowingly departed from utilitarianism in which everything can be compared by one all-encompassing notion of utility because he assumed that there exist more 'worthy' deeds. However, Mill's credit for human's higher needs created the problem. Economists do not like human declarations, because very often they do not coincide with the real choices (Beshears, Choi, Laibson, & Madrian, 2008). For example, Americans ranked highest "talking with children" and 17th "watching television" while it seems clear that in practice most Americans spend much more time watching TV than they ever do talking to their children (Juster, 1985). When people are asked which programs they want to watch on TV, they point out performances, theatre or opera but when it comes to choice, they watch soap operas. However, these results do not mean that people are hypocrites and that their declarations are only the way to look better in the eyes of others. The majority of them do want to have 'worthy' life, not the life that is only based on simple pleasures. However, the existence of weakness of will results in the situations when they do not achieve what they really want.

4.6. Conclusions

The philosophical discussion on weakness of will was necessary to demonstrate a distinction between desires and intentions. Thanks to precommitment strategies, we can explore what people's real preferences are. This knowledge is essential in the context of choices over time because they cannot be compared only by utility measurements. This is why we need to have values that are independent of utility and intentions can indicate what these values are. We need to know what values, independent of choice, individuals have to decide which self should be in control.

Switching into the new approach can be problematic for many economists because economic approach with instrumental rationality and *ad libitum* utility

work perfectly because its assumptions are always fulfilled (tautological system). The weaknesses and controversies connected with this approach were indicated in Chapter 2 in the reference to Becker. In this section, some of these problems were developed from the perspective of choices over time. Moreover, in the context of the weakness of will, one could see how similar some positions in philosophy can be to economics. As it was indicated before, many economists believe in tautological assumptions of their theory (Thaler, 2015) and because of that, they judge values as something unscientific. However, this logical positivistic view of science does not stop there. Many economists believe that economics with axiomatic assumptions is a positive, hard science that creates universal rules (Blaug, 1992; Caldwell, 1982; Fiedor, & Ostapiuk, 2017; Hands, 2001; Hardt, 2013, 2017; Hausman, & McPherson, 2006; Lawson, 1997, 2003; McCloskey, 1983, 1985; Ostapiuk, 2019a; Reiss, 2013). Thus, before introducing values independent of choice it is necessary to have an epistemological discussion and show that neoclassical economists' methodological confidence is unfounded. It will be indicated that ironclad methodologies do not exist anymore, which was demonstrated by the philosophers of science in the second half of the 20th century. However, it will be argued that scientists (economists especially) should not be afraid of relativism, because it is possible to have a discussion on values, and science can thrive in the relativistic world. The acceptance of methodological pluralism and the relativistic nature of science will be used to justify the existence of normative values in economics. Thanks to that, deontological system can be presented as an alternative to the consequentialist system used in economics.

Chapter 5

Philosophy of science and value-free economics

*All epistemology begins in fear – fear that the world
is too labyrinthine to be threaded by reason;
fear that the senses are too feeble and the intellect too frail;
fear that memory fades, even between adjacent steps
of a mathematical demonstration; fear that authority and convention blind;
fear that God may keep secrets or demons deceive*

(Daston, & Galison, 2007, p. 372)

5.1. Introduction

It is neither possible nor necessary from the perspective of the main goals of this book and specifically of this chapter to present the main ideas which have been developed over centuries in the philosophy of science. The concepts highlighted in this section are very selective. They do not reflect the sophistication and depths of debate in the philosophy of science. The purpose of the introduction is to point out that the goal of enlightenment endeavor of science (finding universal laws and ultimate Truth) has not been achieved, and everything indicates that it will not be achieved in the future. Although this statement is well known in philosophy, many economists still believe in the positivistic foundation of objective science. What is more, it seems that many economists agree with the statement that 'economics is what economists do'¹ and they dismiss the methodological discussion about their field. This situation is not exclusive to economists. Also, representatives of the other scientific disciplines treat philosophy of science as something unnecessary. This attitude is attributed to Feynman who wrote that "philosophy of science is about as useful to scientists as ornithology is to birds." The analysis of the insights from contemporary philosophy of science should show that economists' belief in logical positivism and Truth is unjustified. The rules of the scientific game have changed and economists cannot work as they used to.

¹ The phrase 'economics is what economists do' has long been attributed to Jacob Viner (e.g. (Spiegel, 1987, p. 814)) but its origin has not been established to this day (Backhouse, Middleton, & Tribe, 1997).

The quotation which opens this chapter indicates that all epistemology started with the fear of relativism. People want to know what is the Truth, what is right and what is wrong. It is visible in the case of small children who ask questions about (literally) everything. They want to establish some rules by which the world and society are governed. Even our language is constructed to create the dichotomy between truth and false. When one makes a statement 'I eat scrambled eggs for breakfast' it can be only true or false which assumes that something as the Truth exists. Human brain is constructed to find causation of some events and to find patterns everywhere. Therefore, it comes to people as something natural that the goal of science should be finding patterns and the Truth about reality. However, we do not want just to know something. We want to know for sure. We want to have knowledge (justified true belief). Consider the situation when someone points out a few numbers that one believes guarantee winning the lottery and one will win. Then, one has the belief that occurs to be true. However, the last ingredient of knowledge, namely justification, is lacking. In science, we do not want to have just accidental laws (true by accident), but scientific laws (we know why they occurred, or at least we can predict future).

One of the main purposes of philosophy was to find knowledge. This endeavor started to be essential in the Enlightenment. To this day, people who are not familiar with the philosophy of science perceive knowledge as something objective which is independent of subjective perception. Moreover, they believe in universal laws and truths which are waiting to be discovered. This strong belief in science is now pejoratively called scientism and after the 1960s, not many (if any) philosophers share this view.

Until the 1960s logical positivism with a belief in objective science was very popular among scientists. However, from the 1970s on it started to be not only obsolete but something that philosophers sneer at. The term 'positivism' is now exclusively a "philosophical Booadjective" (Hutchison, 1981, p. 204). Despite the criticism, later in this section it is demonstrated that some ideas of this movement are still alive in economics. Popper was one of the philosophers who make logical positivism obsolete to some degree. He is admired by many scientists, which is an exception, because philosophers generally are not praised by scientists. Popper is especially respected by economists, but for the wrong reasons (it will be explained later). Why Poppers is so appreciated? He was a critic of logical positivism, because he believed that science is based on deduction, not induction. First comes hypothesis and later we check it empirically. Not the other way around. For example, if we flip a coin to see what falls out and if we want to repeat this experiment and make it comparable, we will need to assume that we will use the same type of a coin or that the weather conditions will be the same. Popper gives the example of his students to show why theory precedes observation. During the lecture, he asks them to observe

and write down the observation. The students immediately ask “Observe. But what?” Popper points out that our observations are theory-laden to show that it is essential for scientists to have a critical attitude towards their theories. As human beings, we tend to look for patterns and theories. It is very easy for scientists to confirm any theory. The pieces of evidence are all over the place and smart people can find many of them to justify their theories. Popper thinks that the fact that scientists can find proofs for their theories does not mean that their theories are true. Scientists need to embrace the idea of fallibilism² and keep in mind that their theories or they can be wrong. What Popper demands from scientists is always to be critical in spirit. The problem of confirmation is not only a psychological tendency. The very method of confirmation is not valid. The problem is called raven’s paradox (see (Hempel, 1945)). In this example observing black raven confirms our hypothesis that all ravens are black. The paradox works as follows: 1) assume that any observations of an *F* (our raven) that is also *G* (black) confirm ‘all *F*’s are *G*’s,’ 2) my shoes are neither black nor a raven, 3) premise (2) confirm the general claim ‘all non-black thing are non-ravens,’ 4) this is logically equivalent to ‘all ravens are black,’ 5) thus observing my shoes confirms ‘all ravens are black,’ 6) this is absurd, 7) premise (1) is false by *reductio ad absurdum*.

Popper proposes falsificationism to resolve this problem. For him a theory is scientific not when we can confirm it, but when we can find empirical proof that dismisses our theory (Popper, 1959). So, in the case of the hypothesis ‘all ravens are black’ one should look for white ravens that can falsify one’s theory. Falsificationism has two purposes: 1) demarcation (distinguish science from non-science), 2) methodological (how science should be practiced) (Hands, 1993). It is essential to underline that the demarcation criterion only demands that it is logically possible to falsify the theory. It does not imply that we conduct such a falsification (we do not need to find a white raven). Popper thought that a demarcation line between science and non-science is essential, because he observed how different theories could be. Freudism and Marxism were utterly different from Einstein’s theory of relativity. In Freudism and Marxism, it was possible to explain two opposite pieces of evidence by the same theory. In the first case, having a strict father could explain that the child grew up to be introvert as well as extrovert. In the second case, the lack of predicted revolution could be explained by changes in means of production. For Popper, theory that explains everything explains nothing. However, it does not mean that these theories are meaningless.³ They can be insightful and true, but they are not scientific.

² In a sense that no beliefs (theory, view, thesis, etc.) can be conclusively supported or justified (Hetherington, n.d.). This is why scientists always should be critical in spirit.

³ Because of this view Popper restored in some way metaphysics which was abandoned by logical positivism.

For Popper, Einstein's theory of relativity is the best example of a scientific theory. Einstein proposed the bold hypothesis which could be easily falsified because of the risky prediction (very specific). What is most important for Popper is that even if a theory is not falsified, it does not mean that it is true. Acceptance of some theory is always only provisional and can be falsified in the future by a new experiment.

Falsification is right in spirit, but it has its own problems. Because every theory can occur to be wrong in the future, none of them is absolutely true. If a theory successfully passes falsification, the only thing that can be said is that this theory is closer to the truth. This is why Popper proposes the concept of 'verisimilitude' which should indicate which theory is closer to the truth. Thanks to verisimilitude it should be possible to decide which theory is better (that which was falsified 1000 times or that was falsified only one time). However, this method had never worked and the comparisons could not be made. The criticism started with Tichý (1974) and Miller (1974). Oddie (1986, p. 164) claims that Popperian literature has produced "embarrassing results". Many Popperians were also critical. Agassi (1988, p. 473) named it "a boo boo". Although Popper has never given up on the concept verisimilitude, he did ultimately describe his attempt as "admitted failure" (Popper, 1983, p. xxxv).

Quine is the next philosopher in the story. He is considered to be one of the most influential philosophers of the 20th century. He is crucial in our story, because he criticized both logical positivism and falsificationism. Quine (1951) famously rejects the analytic/synthetic dichotomy used by the logical positivists. For Quine due to the 'web of belief,' there is no epistemological way to make a thick distinction between logical and empirical dimension. His argumentation involves language and it is too large topic to be covered here. From the perspective of this chapter, a more important contribution of Quine is theory of underdetermination. In a nutshell, this means that even if we refute some scientific theory by empirical evidence, these tests are not definitive, because no theory is ever tested in isolation (Duhem-Quine thesis, see (Harding, 1975)). To conduct the empirical test many auxiliary hypotheses must be assumed. For example, the testing technique, the boundary conditions, the values of constants, the role of *ceteris paribus* and many others (Hands, 2001, p. 108). Let us get back to the raven example. Before we make the test we need to know what kind of a bird a raven is because there can exist some white birds similar to raven or that "in the process of spray-painting the local bridge the painters accidentally sprayed a raven nest, as to suggest that there really are nonblack ravens" (Hands, 2001, p. 108). Because of underdetermination when contradictory evidence is discovered, there is no way to know whether auxiliary hypothesis or theory itself is wrong. Sometimes it is easy to find out what is the problem with theory. For example, if one conducts an experiment that shows

that copper does not conduct electricity, it does not mean that one falsified the well know theory. Rather it means that someone messed the test up. However, often it is impossible to know where lies the problem with some theory. The discovery of the planet Neptune is the example of this problem. When it was found that the motion of Uran did not match the predictions of Newton's law, the scientists did not reject Newton's laws, but the theory about seven planets in solar system. They found the eighth planet (Pluto). In the end, it means that scientists need to choose one option without a conclusive argument.

So, if we do not have any rational rules to compare different theories, how we should decide which is better? Kuhn has the answer to this question, but his answer generally undermines the rational view of science. His book *The structure of scientific revolutions* (1962) is probably the most known book in the philosophy of science. In addition to philosophers, it draws the attention of psychologists and sociologists.⁴ Kuhn underlines that every theory is theory-laden and he introduces the idea of 'scientific paradigm' which can be explained as the lens by which scientists look at reality. Kuhn argues that scientists cannot rationally decide which theory is better and they choose, as a community, between paradigms. For example, in the 16th century scientists supported the Copernican system and the Ptolemaic system started to be obsolete even though at that time the evidence did not support the Copernican theory. Kuhn as a historian of science observed that scientists do not treat anomalies as counterexamples of their theories. In the 19th century, scientists did not dismiss Newton laws, because the predictions did not match. Instead, they changed the auxiliary hypothesis and they searched for a new planet. This example shows that scientists do not always follow falsificationism and it can be beneficial for science (discovering Neptune). Therefore, paradigm means that scientists just trust in some theory and they look at reality through the lens that some theory provides. Often paradigms are explained as 'gestalt shift' when one can observe the pictures of rabbit/duck or old women/young lady. It is only possible to see one character at a time and it depends on the perspective from which we look at these pictures. Kuhn argues that the different paradigms are incommensurable with each other. We cannot rationally compare them, because they have completely different languages by which they describe reality. So-called facts will not help scientists

⁴ Many ideas displayed in his book had been presented before by Fleck in *Genesis and development of scientific fact* (2012) which had been published 30 years before Kuhn in German (1935). From the sociological perspective one can see why it were the ideas of Kuhn that have gained popularity. After the Second World War people were reluctant toward science. They had in mind totalitarian systems, ecologic disasters and atomic bomb. People discredited the idea that science could provide well-being for everyone. The 1960s were the times of counterculture which looked for the alternative ways of living in opposition to science. It created the demand for philosophy that could dethrone science from the epistemological pedestal (Sady, 2013, p. 144).

with the decision, because how facts are interpreted depends on the paradigm that they believe.⁵ The last crisis is the epitome of this problem. The left argues that crisis was caused by lack of regulation, whereas the right argues that there were too many regulations. In the end, it means that no amount of evidence can decide between paradigms. What is essential for Kuhn, is the fact that paradigm does not concern only one person but the whole scientific community. This is why he underlines the context of discovery (science is embedded in society).

Kuhn argues against the enlightenment view of science where knowledge cumulates with time. Kuhn (1962) claims that science develops from revolution to revolution which he calls “paradigm shifts.” Before the revolution, in time of ‘normal science’ scientists follow paradigm and they are dogmatic. They try to dismiss anomalies and find a way to defend their paradigm. However, with time a paradigm has to deal with more and more anomalies. When a paradigm is unable to deal with anomalies and cannot resolve accumulating problems, then science will fall into crisis. The dogmatic scientists start to be open for a change and they are more open for philosophical reflection on their work. Scientists want to know some new rules that would make their work successful again. We could observe this mechanism in the 1980s when economists started to be interested in philosophy and philosophy of economics emerged. Scientific revolution is similar to social revolution, because in both cases there are no rational arguments to decide which theory is better. This is why, in the end, the scientific community decides which paradigm is better. Very often this process takes the form of a struggle between older scientists who try to defend their theory and young scientists who criticize old order.⁶ In the end, the group which wins is the one which has more power, more persuasion and more supporters (Sady, 2013, p. 163). In economics, we could observe this struggle in the 1980s when behavioural economics criticized the neoclassical paradigm. At the beginning, the conclusions from behavioural economics were in many ways dismissed (Thaler, 2015). However, recently behavioural economics started to be identified with the mainstream (Angner, 2019; Thaler, 2016; Truc, 2018).

Kuhn has changed the game for the philosophers of science. From that time, they have had to agree on the relativity of the scientific method. However, scientists have been trying to escape from absolute normativity by creating rules that can govern the scientific method. Lakatos is one of them. In Chapter 2, the ‘scientific research programme’ conception was analyzed. Therefore, here Lakatos is presented from a different angle. In the context of Kuhn and Popper. In general, Lakatos uses insights from both Popper and Kuhn, but in many cases, he dis-

⁵ In economics Kaldor (1957) writes about “stylized facts.”

⁶ The German physicist Max Planck even wrote that “a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it” (Planck, 1950, p. 33).

agrees with them. Although he assumes that ‘hard core’ does not require to be falsified, he emphasized the importance of empirical content which is thoroughly Popperian. Lakatos, like Popper, describes the empirical content of a theory to be “the set of its potential falsifiers: the set of those observational propositions which may disprove it” (Lakatos, 1980, p. 14). Therefore, empirical data provided by the auxiliary hypothesis should be falsifiable. Secondly, Lakatos highlighted the importance of novel fact with the requirement of “independent testability” (Hands, 2001, p. 112). It means that we cannot avoid any falsifying evidence. For example, we cannot state that “all ravens are black except the white one you just showed me” (Hands, 2001, pp. 124, 125). This kind of strategy is called *ad hocness*, and science needs to rule this out if it wants to have the real progress, as was indicated in the case of Becker’s economic approach.

However, in many cases, Lakatos disagree with Popper and his ideas are more in line with Kuhn. The obvious example is the fact that ‘hard core’ is unfalsifiable. Moreover, Lakatos argues that all theories are “born refuted” (Lakatos, 1980, p. 128) and that the progress comes from the corroboration, not falsification of novel facts (Hands, 2001, p. 205). Lakatos also uses a historical meta-methodology in which theories are assessed from a historical perspective (progressive or not), whereas falsificationism is the method independent of history. The differences from Popper make Lakatos popular in economics, because he enables scientists to choose between different theories.

Although Lakatos agrees that scientists are governed by paradigms, he wants to understand why scientists choose the particular paradigm. Often scientists do not know what criteria they use to assess scientific theory or if they know, they indicate contradictory criteria. Lakatos does not think that scientists are irrational, but that they just do not know the whole picture of their choices. Lakatos thinks that the goal of science is to reconstruct the decisions of scientists about paradigms rationally. He believes that thanks to this process, some universal criteria of choice between paradigms can be found (Sady, 2013, p. 183). From Lakatos’ perspective, some research programmes can be assessed only with hindsight. Scientists need to wait and see if each programme is progressive or not. Lakatos writes about three concepts of progress. The first one, “theoretical progress” means that new theory “has some excess empirical content over its predecessor, that is, if it predicts some novel, hitherto unexpected fact” (Lakatos, 1978, p. 33). The second one, “empirical progress” takes place when “some of this excess content is also corroborated, that is, if each new theory leads us to the actual discovery of some new fact” (Lakatos, 1980, p. 34). Lakatos describes a research program as a “progressive,” if it is “both theoretically and empirically progressive, and degenerating if it is not” (Lakatos, 1980, p. 34). Finally, the third concept – “heuristic progress” – specifies that the changes in protective belt are consistent with the hard core of the programme. Lakatos’ notions of

theoretical and empirical progress presuppose that the changes in question are in accordance with heuristic progress (Hands, 2001, p. 204). For Lakatos, the scientific progress takes place when degenerating research programme is replaced by a new progressive one. However, each scientific programme can change both from progressive to degenerative and vice versa.

The last philosopher in the story is Feyerabend (1993). He introduces the idea of epistemological anarchism which is captured by the phrase ‘anything goes.’ Sometimes scientists misinterpret his work and perceive him as a some kind of a lunatic. However, we need to remember that he was influenced by Nietzsche and in some way he purposefully provoked scientific community (Feyerabend, 2005, p. 48). Moreover, he did not criticize science for the sake of criticism. He displayed some problems of science because he wanted science to do better. Firstly, he argued against authoritative and dogmatic science which behaved like a church. Feyerabend cautioned against people who thought that they knew what the Truth was. In this sense, he was a fierce critic of fundamentalist idea of the Truth. For him science started to behave like church because it believes in its own infallibility and tries to make scientific method the only proved tools for understanding the world. Because of this criticism, some can think that Feyerabend was against science as a whole. It must be stated that he fought with fundamentalist science, not science in general. He would not state that ‘anything goes’ and that vaccine deniers would be right to hold their views.

Secondly, he argued that science should not have any methodological rules. Because none of them guarantee success. He challenged scientists to present him some scientific methodology and he would prove that it did not work at some point in history. Feyerabend (1975) enjoyed giving examples of theories that were discovered precisely because scientists did not follow the ruling methodology at that time. He showed that a particular methodology could be effective in one case and ineffective in another. It means that scientists just do not know which rules work and which do not. Therefore, we all should be anarchists in methodology because it can lead to new discoveries in science. It comes as no surprise that Feyerabend criticized every methodology (e.g., induction, deduction, Lakatos, falsificationism). Although he criticized falsificationism, he was Popperian in spirit (Feyerabend, 1961), for a discussion, see (Tambolo, 2015; Collodel, 2016). For Feyerabend scientists constantly had to criticize their theories. They should not do this when their theory turned out to be wrong and falsified, but much earlier, when the theory had been successful. This is the only way to defend oneself from dogmatism. Even though Feyerabend admired Lakatos,⁷ he argued that knowledge about the fact that paradigm was progressive or degenerative was useless as a tool to decide between theories. It is a historical perspective and

⁷ He wanted to write book with him but premature death of Lakatos stood in the way.

theories in the future can switch from degenerative to progressive programmes and vice versa. Moreover, Feyerabend (1993, p. 163) argued that Lakatos' scientific research programme lead to relativism, because the only way for a scientific community to support progressive and to abandon degenerative programme was by institutional violence (withdraw money for research that deals with the degenerative programme, do not publish their work, etc.).

Of course, the history of philosophy of science has not ended on Feyerabend and 1980s. From that time on many different approaches in the framework of constructivism have been developing (e.g., feminist epistemology, the sociology of science, postmodernism, and performativity); see these approaches in the context of economics (Boldyrev, & Svetlova, 2016; Brisset, 2016; Callon, 2006; Mackenzie, 2006; MacKenzie, Muniesa, & Siu, 2007; Scheuer, 2015; Zboroń, 2009a, 2009b, 2013). Moreover, naturalism started to be popular among scientists over the last years (e.g., (Beed, & Beed, 2000; Caplin, & Schotter, 2008; Davis, 2016; Dekker, & Remic, 2018; Ross, 2005)). All of these approaches are significant, but due to limits of this book the discussion must end right here. The essential fact about all these approaches is that they use the insights from Kuhn, Lakatos, and Feyerabend. All of them criticize fundamentalist epistemology of science. They underline the problems of theory-ladenness, underdetermination, socially constructed knowledge, representation of reality, objectivity, and scientific facts.

The purpose of the discussion on the philosophy of science was to present that fundamentalistic idea of epistemology in which there is objective Truth is no longer an available option. There are no firm rules by which sciences can be cultivated. However, many economists do not care about this state of affairs and they do science as if nothing happened. To this day, many economists are positivistic in spirit and they believe in facts, falsificationism, independence between theory and reality, and is/ought distinction (Fiedor, & Ostapiuk, 2017; Gorazda, Hardt, & Kwarcinski, 2016; Hands, 2001; Hardt, 2017; Hausman et al., 2017; Kincaid, & Ross, 2009; Lawson, 2015; Mirowski, 2002; Putnam, & Walsh, 2011; Reiss, 2013; Rodrik, 2015).

5.2. Positivist methodology of economics

Over 30 years ago McCloskey described the methodology of mainstream economics, and in many ways this description is still accurate. Before starting our analysis, some arguments have to be made to justify McCloskey' almost 40-year-old description of methodology of economics. As was indicated at the beginning of my book, many scholars argue that recently economics has transformed due to the incorporation of the knowledge from other social sciences like psychology and sociology. Therefore, contemporary mainstream economics is different from orthodox neoclassical economics of the 1950s (e.g., (Cedrini, & Fontana, 2018;

Colander, 2000; Colander et al., 2004; Davis, 2006)). The paragon of economics' transformation is behavioural economics which started to be perceived as a new mainstream (Angner, 2019; Thaler, 2016). Moreover, it is argued that economics is not so theoretical anymore, and we could observe 'empirical turn' in the last years (Backhouse, & Cherrier, 2017). In general, the process of opening economics to other social sciences is called 'reverse imperialism.'

However, many argue that despite the changes, the foundations of economics have not changed and mainstream economics works as it used to (Berg, & Gigerenzer, 2010; Dow, 2012; Lawson, 2015; Madra, 2016; Ostapiuk, 2019a). Everyone agrees that economics has incorporated many insights from other social sciences. However, they are perceived only as adjustments, whereas the paradigm has not changed. Many economists still rely on positivism and believe in objective and value-free economics which is based on deduction and axiomatic revealed preference theory (viz. rationality, utility, and *homo economicus*).

Therefore, McCloskey's 40-years-old description of the methodology of economics is, in many ways, still accurate and worth analysing. She highlights (1983, pp. 484, 485) its main precepts such as:

- (1) Prediction (and control) is the goal of science.
- (2) Only the observable implications (or predictions) of a theory matter to its truth.
- (3) Observability entails objective, reproducible experiments.
- (4) If (and only if) an experimental implication of a theory proves false is the theory proved false.
- (5) Objectivity is to be treasured; subjective 'observation' (introspection) is not scientific knowledge.
- (6) Kelvin's Dictum: 'When you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.'
- (7) Introspection, metaphysical belief, aesthetics, and the like may well figure in the discovery of an hypothesis but cannot figure in its justification.
- (8) It is the business of methodology to demarcate scientific reasoning from non-scientific, positive from normative.
- (9) A scientific explanation of an event brings the event under a covering law.
- (10) Scientists, for instance economic scientists, have nothing to say as scientists about values, whether of morality or art.
- (11) Hume's Fork.

Due to the fact that in the first chapter the problems with value-free economics were indicated, in this section the influence of logical positivists on economics is presented mainly in the context of the mathematization of economics. This process which sometimes is called "formalist revolution" (Blaug, 2003) is one of the reasons why it is possible for economists to still hold neopositivistic view of science even though it is obsolete. Economists wanted economics to be like

hard science (physic envy). In order to have predictive power, they search for universal rules and models which are unrealistic.⁸ Many papers, and books were written about this issue (Beed, & Kane, 1991; Debreu, 1991; Godłowski-Legiędź, 2014; Mirowski, 2002; Ostapiuk, 2017a, 2019a; Weintraub, 2002) and this topic is well-analysed. Therefore, only the most important issues about mathematization of economics are underlined in this section. Many economists follow Samuelson who writes that “mathematics is language” (Samuelson, 1952, p. 52). However, it is important to understand that economists did not invent this idea. They took it from the only philosophy they knew, at the time, logical positivism (Mirowski, 2004). Around that time, Carnap was advertising a vision of philosophy as “the mathematics and physics of language” (Richardson, 2003, p. 180) and had just published his *Logical syntax of language* (1937) where he tried to provide a metalanguage for all of science. Mathematical language is crucial for neoclassical economics because it provides economists with a tautological system in which every theory is consistent within itself. The assumptions of economic approach or welfare (revealed preference) are always fulfilled thanks to that. Many assumptions of neoclassical economics are purposefully unrealistic (e.g., representative economic agent). Due to that, economists can dismiss descriptive criticism from psychologists or sociologists.⁹

Thanks to tautologicality of the assumptions, the economic approach works. This is why so many economists still believe in value-free economics and they champion mathematical approach towards economics. Vickery (1964, p. 5) indicates that:

Economic theory proper, indeed, is nothing more than a system of logical relations between certain sets of assumptions and the conclusions derived from them. The propositions of economic theory are derived by logical reasoning from these assumptions in exactly the same way as the theorems of geometry are derived from the axioms upon which the system is built.

After the Second World War some economists started to look like the scientific community in *The glass bead game* (Hesse, 2000) where scientists live in an ivory tower and they do not care about reality. Blaug makes this point when he writes that “economy increasingly become an intellectual game played for its own sake and not for its practical consequence for understanding the economic world” ((Blaug, 1997, p. 3), similar point was made by Lipsey (2001)). Many economists create theories for theory’s sake. It does not matter if they are descriptively true. The most important thing is if these theories are closed axiomatic systems that work. This fascination with the effectiveness of theories, not with their

⁸ Using models is something that distinguish economics from the other social sciences, see (Leijonhufvud, 1973).

⁹ The role of idealistic models in economics, see more in: (Hardt, 2013, 2016, 2017; Mäki, 2002, 2006; Morgan, 2012; Reiss, 2013; Sugden, 2000, 2009b).

accuracy towards reality, displays economics' schizophrenia. On the one hand, economists are naïve realists who believe in objective (measurable) data and logical positivism. On the other hand, they use instrumentalism in which the truth cannot be known, and prediction is the only thing that matters because of its usefulness. This inconsistency is the result of a lack of methodological knowledge and the opportunism of economists. They use methodologies that suit them. Objective data give them a rigid and scientific foundation that differs them from the other social scientists. On the other hand, instrumentalism with disbelief in the Truth gives economists justification for the simplified mathematical models which do not need to explain reality. They only need to have predictive power.

The arguments against objective science were presented earlier. Therefore, in this section they are investigated only in the context of falsificationism which is preached in neoclassical economics. There are two main problems associated with falsificationism in economics. Firstly, economists never really have practiced it. Blaug makes this point when he writes that "modern economists frequently preach falsificationism [...] but they rarely practice it: their working philosophy of science is aptly described as 'innocuous falsificationism'"¹⁰ (Blaug, 1992, p. 111). Of course, economists engage in empirical research which should make their theories more prone to falsification (it is not a metaphysics in the end, but facts). However, as Blaug (1992, p. 241) shows:

much of it is like playing tennis with the net down: instead of attempting to refute testable predictions, modern economists all too frequently are satisfied to demonstrate that the real world conforms to their predictions, thus replacing falsificationism, which is difficult, with verification, which is easy.

Therefore, in the opinion of Blaug, economists do not really falsify their theories but rather verify them, which shows the lasting influence of logical positivism on economics.

The problems with falsificationism are deeper. Many economists have insufficient knowledge about underdetermination and theory-ladenness. They see falsificationism as a special type of empiricist foundationalism (logical positivism) but without the problem of induction (Hands, 2001, p. 292). Because Popper was never an empirical foundationalist, it means that economists are Popperians for the wrong reasons (Hands, 2001, p. 292). Secondly, if they read Popper in a modern way, they would not be Popperians, because falsificationism does not provide a methodology that demonstrates how science should be done. Moreover, it does not establish the firm demarcation line between sciences and pseudo-sciences. In the end, one cannot escape from metaphysics.¹¹ Hands

¹⁰ Blaug owes this formulation to Coddington (1975, p. 542).

¹¹ Lawson (1997, 2003) shows the failure of dismissing metaphysics in neoclassical economics.

(2001, p. 293) writes that “the falsificationist methodology of bold conjecture and severe test provided a set of rules for the game of science without providing an ultimate aim or purpose for playing the game”. However, the goal must be provided which was indicated by Lakatos (1980 p. 154): “the rules of the game, the methodology, stand on their own feet; but these feet dangle in the air without philosophical support”.

Now, it is time to present some criticism of instrumentalism which is popular methodological stance in economics. Its main premises were presented in the second chapter in the context of Becker’s economic approach. To show a problem with instrumentalism, the philosophers of science use the example of Ptolemy’s astronomy where it is assumed that the Earth is at the centre of the Universe. Although the theory is false, it has strong predictive power. The theory was useful, because it guided a navigation and helped in the creation of a more accurate calendar. If judged only by its usefulness, we would hold this theory until now. However, it is obvious that nobody treats Ptolemy’s astronomy seriously despite its usefulness.

In general, instrumentalism is not wrong in stating that there is no Truth and this is why the ultimate goal of science is usefulness. Instrumentalism has its roots in pragmatism but seems to go too far. Pragmatists like William James or Charles Peirce did not argue against the possibility to decide if theories are true or false. They just argued that theories, in the end, must serve practical ends (‘cash value’) (Capps, 2019; James, 1981). Cash value of a theory is not only its predictive power. If theories are like instruments that we use, we need to know for what purpose we need them. A barometer measures air pressure and because of that people can predict and anticipate the changes in the weather. However, other tools are necessary to control the weather. A barometer alone can only assist in measurement or observation (Hausman, 1998, p. 190). Conversely, a hammer helps one to act. Both these instruments have different purposes. Therefore, the idea of theories as tools is ambiguous, which was pointed by Hausman (1998, p. 190): “Should it aim to give them tools to measure and passively to predict, or should it aim to give them tools to act and to control?”.

We should step down from these abstractions towards real economic problems. For a while, Phillip’s curve enabled economists to predict the rate of unemployment connected with the rate of inflation. However, the relationship ended when the governments started to use this relation to regulate the rate of unemployment. Why did it happen? Prediction requires only knowledge of stable correlations, whereas control demands knowledge of causation (Hausman, 1998, p. 190). Instrumentalists do not care about this distinction, because they believe that the only practical goal is a prediction. They hold this view because they believe that it is impossible to know the Truth. As a consequence, they did not notice that the knowledge of causes had a practical end (control). This is why Hausman (1998)

argues that search for causes is central to science. In general, most philosophers of science are not instrumentalists. Most of them agree that the goal of science is both predictive and explanatory (Cartwright, 1989; Gorazda, Hardt, & Kwarciński, 2016; Hands, 2001; Hausman, 1992; Reiss, 2013).

The criticism mentioned above has its bearing on value-free economics. Firstly, we know that prediction is not the only thing that matters and that explanation is also important. This is why economists cannot create theories whose only purpose is to work logically. Economics cannot be only an intellectual game. Secondly, as was indicated in the first chapter, there is no fundamental and unbridgeable distinction between normative and positive approaches. We also know that the economic theory of well-being (welfare) contains normative statements. From this perspective, it seems justified to argue that economics should be closer to reality and help people in achieving a good life. To do that, economists need to consider normative approaches as full-fledged scientific programmes. Many economists still believe in is/ought distinction and they consider a discussion about values as something meaningless. As a scientific community, we need to understand and accept the fact that maybe we will never find the conclusive answers about people' ends. However, it does not mean that we cannot, and should not discuss them.

5.3. Weber and value-free economics

It has already been presented where value-free economics leads. Paradoxically, Max Weber' works show why value-free economics faces unresolvable problems and why economists need to investigate its values. It can sound paradoxical, because many economists perceive Weber as a founder of value-free science (Hammersley, 2017; Klimczak, 2014; Mongin, 2006; Van Dun, 1986). Many economists used his work to support the is/ought distinction. This misinterpretation of Weber is possible, because Weber's ideas are complex and broad, which can lead to ambiguity (Weber, 1949). There is much discussion whether by value-freeness he means ethical neutrality of researcher, i.e., we should deal with facts without an axiomatic ingredient, or that we should not express normative statements at all (Klimczak, 2014).

To begin with, it is essential to underline that Weber did not try to present a systematic epistemological treatise which would end the 'strife over methods' (*Methodenstreit*) of his time between positivism and historicism. His goal was much more pragmatic. Wolin (1981, p. 414) writes that Weber "formulated the idea of methodology to serve, not simply as a guide to investigation but as a moral practice and a mode of political action". For Weber ethics and science go hand in hand. He demands intellectual honesty from scientists who make claims about scientific truths, hypotheses, and meaning postulates (Van Dun, 1986,

p. 19). Weber realizes that scientists cannot escape from values, so they should scrutinize them honestly. From this perspective “it would perhaps be better to drop the term *Wertfreiheit* altogether, and to speak only of ‘freedom from prejudice’” (Van Dun, 1986, p. 20). The only place where Weber is explicitly against values in science is when he argues against the lecturers who impose some ideologies on students (Kalberg, 2005, p. 318; Weber, 1919).

In Weber’s opinion, scientists cannot escape from values in their scientific work. Even one’s interest in some field is triggered by some values that one cherishes. Because in many ways Weber was pragmatic, he understood that if scientists cannot eliminate their values, they should at least scrutinize them rationally. Moreover, this assessment will provide a great benefit for science. To scrutinize values, Weber makes a distinction between ‘instrumental value judgments’ and ‘methodological value judgments.’ ‘Instrumental value judgments’ should make a scientific discussion possible and fruitful. They can take the form of: (1) assessing the internal consistency of the value premises from which various normative judgments are derived, (2) deducing the implications of these value premises in the context of the specific circumstances to which they are used, (3) observing the factual consequences of different ways of achieving normative judgments (Runciman, 1968, pp. 564, 565; Weber, 1949, pp. 20, 21). ‘Methodological value judgments’ cannot be checked, but without them science could not exist. They are the main assumptions of scientific approaches. In value-free economics ‘methodological value judgments’ are for example invisible hand, equilibrium and effectiveness of markets, instrumental rationality (WARP), and revealed preference theory. Even though we cannot find an objective basis to compare them, we need to investigate these values. Scientists must practice a “contemplative attitude” towards these values (Schweitzer, 1970, pp. 18, 19). Contemplation here does not mean indifference towards values but impartiality. Scientists must accept subjectivity of science, and this is why they should be self-consciously subjective. What kind of implication it has on value-free economics? Most importantly, many economists tried to eradicate every value from economics. In the light of Weber’s distinction, it is impossible, because we still have ‘methodological value judgments’ in economics. Economists should use Weber’s insight, because it enables them to realize the normativity of their fundamental assumptions which many treated as positive. Thanks to that, they will be better prepared to see the implications of their theories. Moreover, it would encourage them to dismiss the assumptions that do not work.

The concept of ideal type is the next methodological idea of Weber that is essential in the context of value-free economics. ‘Ideal type’ is a methodological construction which is based on some fictional assumptions about people or reality. We exaggerate and emphasize significant characteristic of social phenomena and then we classify them into a model that is useful for the

researcher (Reiss, 2013, p. 110). It goes without saying that ideal type is only simplification and its purpose is not to describe people's behavior correctly.¹² In economics, there is a long tradition of using ideal types. Machlup (1978, p. 298) suggests that *homo economicus* should be perceived as a Weberian ideal type (see also (Morgan, 2006)). The main purpose of ideal type is to predict the future. It should be a useful heuristic device for scientists. However, Weber does not claim that scientists should be solely focused on prediction. He also underlines the importance of generalization about a social phenomenon which has explanatory power (Reiss, 2013, p. 110). Secondly, Weber warned against perceiving ideal types as covering laws that claim universal applicability. For him this misunderstanding has led to "naturalistic monism in economics" (Weber, 1949, p. 86). His warnings come true and some economists believe that *homo economicus* is a real human (Thaler, 2015).

In times of Weber the distinction between is/ought was treated as something obvious and Weber was not an exception. He believed that a distinction between fact and values exists. Nowadays, however, due to Kuhn and others (e.g., (Feyerabend, 1993; Kuhn, 1962; Lakatos, 1980; Rorty, 1979)), we know that this distinction is fragile and even facts are theory-laden in some way. Many scientists tried to reconcile scientific statements with their unavoidable relativisms. Nagel was a philosopher who tried to protect Hume's guillotine. He made a distinction between 'characterizing value judgments' and 'appraising value judgments,' which is similar to Weber's distinction between 'instrumental value judgments' and 'methodological value judgments.' Nagel's 'characterizing value judgments' contain criteria that can be assessed. For example, it can be done by logic or data. Different judgments can be compared thanks to objective criteria. On the other hand, 'appraising value judgments' evaluate states of the world. For example, the desirability of a particular kind of human behaviour or assessment of the outcomes that are created by that behaviour. Scientists cannot choose between 'appraising value judgments' because there are no objective criteria to do so. There is only political and social power. Nagel understands that science cannot work without values. However, he believes that science can be free, at least, from 'appraising value judgments' and scientists can use 'characterizing value judgments' that can be compared by the scientific method (Nagel, 1961, pp. 492–495).

A similar distinction was made by Sen. He distinguishes 'basic' and 'nonbasic' value judgments. Sen explains that "a value judgment can be called 'basic' to a person if the judgment is supposed to apply under all conceivable

¹² Rodrik (2015, pp. 43, 44) presents Borges' story *On the exactitude of science* to make a point about models in economics. The story is about society that builds maps. With time, they want to describe world more realistically (build ideal map) and in the end they build map in scale 1:1. Rodrik uses this analogy because he wants to show that models must be always simplified to be useful.

circumstances, and it is ‘nonbasic’ otherwise” (Sen, 1970, p. 59). This distinction is essential, because it draws the line between sciences and nonscience. If we have basic or pure value judgments like freedom and we do not consider any other value (e.g., well-being), then it is impossible to have a discussion, because discussion assumes that we can change our opinion due to the arguments that are presented. The essence of science is a discussion within which we can change our values and opinions. Thus, only ‘nonbasic’ judgments make science possible.

Blaug also observes that the methods which try to give scientists the tools to discuss values are not ironclad. He writes that “the difference between the methods of reaching agreement on methodological judgments and value judgments is one of degree, not of kind” (Blaug, 1992, p. 116). Blaug (1992, pp. 114, 115) notices that there is a difference between is and ought statement, but “at the deep philosophical level, this distinction is perhaps misleading”. He underlines that so-called facts also must rely on the acceptance of some rule of the scientific game which strives for objectivity. However, it will never be achieved completely, because acceptance of scientific rules depends on the scientific community and it is done by persuasion. The similar point was made by McCloskey (1983, 1985). Although Blaug agrees that science in the end is theory-laden, he thinks that Nagel’s distinction between characterizing (methodological) and appraising (normative) judgments is real and significant (Blaug, 1992, pp. 114, 115). Without this difference, we fall into absolute relativism which does not offer a lot. Some difference between is/ought is necessary, which can be observed in the case of Myrdal (1970) who argues that every statement is equally value-laden.¹³ If we look from black and white perspective, we will not see a grey area in the case of value-ladenness. Of course, the inflation rate is theory-laden. Economists decide normatively which products and services should be inclined in this measurement. However, when the inflation rate is presented, it is the same rate for everyone and this kind of objectivity is essential for science.

The purpose of this subchapter was to demonstrate that scientists cannot escape from values. They always make some ontological, epistemological, and ethical preassumptions. Scientists need to deal with them, even though (especially) it is impossible to decide which of them are true. From this perspective, it seems convincing that economists should work on normative approaches which concern human happiness. Of course, at the core they deal with normative statements. However, thanks to instrumental value judgments, scientists can check if they are logically consistent and if they are supported by the evidence. In the case of methodological value judgments which cannot be eradicated from science, it can be concluded that it is better to put these values on the table and point them out explicitly rather than hiding these values in disguise as it is mostly done in positive economics (e.g., welfare economics).

¹³ However, some argue that this perception of Myrdal is a simplification, see (Hodgson, 2000).

5.4. Consequentialism versus deontology

There are two main systems that investigate human values: consequentialism and deontology. In Chapter 2, the problems of consequentialism in value-free economics were presented where a qualitative difference between values like altruism and self-interest utility was made. It was demonstrated that utility used by economics can mean everything. In this section, the goal is not to compare the two large ethical systems (consequentialism and deontology), but to present and argue for a difference between intrinsic and instrumental values. Intrinsic values are independent of consequences, and their existence is necessary if we want to choose which self should be in control. Kant's deontological system is presented as an attempt to establish the ultimate intrinsic values.

Before discussing these two ethical systems from the above-mentioned perspective, we need to underline the difference between ethical propositions and preferences. The former are the statements about a rank order of preferences among alternatives and they apply to more than one person, whereas preferences (tastes) apply only to one person. If one says: "I prefer x to y ," this is a personal value judgment or a taste. On the other hand, if one says: " x is better than y ," there is an assumption that one expects others to prefer x to y . This is why a moral proposition is a "common value" (Boulding, 1969, p. 2). Individuals often act from moral motives, and this is why economics needs to examine them. Understanding consequentialism and deontology is also important in the context of human happiness. The former approach is connected with subjective well-being, whereas the latter refers to objective well-being.

In general, deontology and consequentialism are completely contrary to each other and always have fierce apologists. In general, consequentialism determines the value of an act by the goodness of the state of events that they lead to (consequences). Deontological theory assigns intrinsic value to some acts which are independent of their consequences. For example, breaking a promise is wrong itself, no matter what consequences are (Broome, 1992). Of course, as every ethical theory, these two also face criticism. However, from the perspective of value-free economics one problem associated with deontology is essential. If one performs an act, the consequences are interconnected with performing an act. So, if the act is intrinsically good or bad, then the consequence can be also good or bad. As Broome (1992, p. 3) puts it:

When evaluating the consequences of the act, there is nothing to stop our including the value of this consequence along with others. In this way the intrinsic value of the act can be taken into account within an evaluation of the consequences.

Therefore, consequentialism can take into account the intrinsic value of acts which seem to be deontological. For example, imagine the moral dilemma

where one made a promise (Williams, 1973). If one keeps it, five promises made by the other people will be broken. However, if one breaks a promise, these five promises will be kept. The rest of the consequences will be equal (*ceteris paribus*). Should this person break her promise or keep it? One answer is that she should break it because breaking promises is intrinsically bad and breaking the promise in this scenario will bring this bad thing. On the other hand, keeping one promise will bring five promise-breaking which is five times as bad. Therefore, it is better on balance to break one's promise and this is what she should do. However, this answer is in the spirit of consequentialism. Although it takes account of an intrinsic value of promise-keeping, it does it within the context of utilitarian calculus.¹⁴ How intrinsic values can be taken into account by consequentialism was presented in the case of Becker. For him it is possible because economics understands utility *ad libitum* and every value (e.g., good, friendship, love) are put in the "bottomless pit" of utility (Ostapiuk, 2017b, pp. 73, 74). Even though sometimes there are problems with establishing the precise line between deontology and consequentialism, it does not mean that it is impossible to distinguish values from utility. Moreover, one needs to this if one wants to construct a comprehensive theory of well-being.

The above-mentioned distinction is necessary because, as was indicated in Chapter 3, subjective well-being approach has many limitations. This utilitarian approach is consequential, because it measures how much happiness each action provides and it assumes that people should behave to maximize utility. For utilitarians, good is defined in pleasure or happiness. There is nothing controversial in saying that happiness has an intrinsic value. However, as was indicated in the case of Freud and drugs, it is problematic to agree that happiness is the only thing that is intrinsically good because freedom has an intrinsic value as well. In the case of preferences, utilitarianism has a problem with answering the question "is the satisfaction of preferences an intrinsic good, like pleasure or happiness, or is it good instrumentally because it leads to consequences that are good in themselves?" (Sagoff, 1986, p. 3). We cannot argue that the satisfaction of preference is intrinsically good, because some preferences are racist, sadistic or unjust.¹⁵ This is the reason why utilitarian calculus cannot be used in some cases. However, to know the limits of utilitarian calculus, we need to establish some values that are independent of utility.

¹⁴ Nozick (1974) famously uses the structure of this example. There is an act (initiation of police force) that violates some people's rights (people will be taxed and Nozick considers it as a violation of rights). However, it will prevent greater violations of rights (murder, rape, robbery, and so on).

¹⁵ Harsanyi (1977, p. 33) rejects the utilitarian calculus in the case of "sadism, resentment or malice". The problem of malicious preferences is frequently used against consequentialist utilitarianism, e.g., malicious dictator that takes pleasure in killing or torturing people.

Value-free economics assumes that satisfaction of preferences is not intrinsically good but it should be instrumentally perceived. So, economists assess the satisfaction of preferences good, insofar as it increases welfare/well-being/utility of the person.¹⁶ However, as was indicated in Chapter 1, this perspective does not mean anything, because economists define utility or welfare in terms of the satisfaction of preferences and they assume that people always maximize their utility. It seems to prove that instrumental reasoning in economics has severe limitations. Moreover, the tautologicality of the single dimension utility that contains everything (every intrinsic value) was observed by many scholars. For example, Brennan (1989, p. 197) writes that “utility is that which an agent maximizes through choices, but as defined within economics it has no empirically or normatively meaningful existence outside those choices”. Because both utility and preferences are defined by choice, it is impossible not to maximize them.

Due to these problems, Brennan is sympathetic to multiple utility framework in which there are different values that cannot be compared in one-dimensional utility framework (Brennan, 1989). As was indicated before, multiple utility framework can be quite useful to distinguish self-interest from altruism (e.g., ‘commitment’ and ‘sympathy’ (Sen, 1977, p. 326)). Lutz agrees with Brennan and argues that multiple utility framework has “heuristic function” because it promotes a better understanding of human behavior (Lutz, 1993, p. 151). Moreover, he indicates five big problems with single-utility maximization. Three of them are essential from the perspective of this book.

1. Because economists assume that agents are unable not to maximize, they cannot see a possibility of market disequilibrium. As Boland (1981, p. 1035) puts it, “those economists who put maximization beyond question cannot ‘see’ any disequilibria”.
2. Single-utility maximization undermines “X-efficiency” (Leibenstein, 1966) and “satisficing” (Simon, 1957). In Leibenstein’s “selective rationality,” people select the extent to which they deviate from maximization behavior. They decide how much to constrain their behaviour for the sake of others. If one insists that “all is maximization.” there would be no place for any degree of “maximization deviation” (Leibenstein, 1979, p. 485). Similarly, Simon’s “satisficing” is meaningful only if it is distinct from maximizing. It is possible when the levels of aspiration are variable. The problem of satisficing was presented previously in the case of instrumental rationality in which satisficing could be explained in terms of maximization. One just needs to add the constraints within which people maximize their utility like time or cognitive abilities.

¹⁶ From mainstream economics’ perspective, preferences ranking is neutral and does not have any intrinsic (good or bad) value.

3. Experimental research in psychology provides evidence of potential costs which are the result of an economic focus on incentives (Lutz, 1993). The subjects of the study engage in different intrinsically meaningful activities like drawing or puzzle solving. In spite of the fact that no (extrinsic) reward is necessary for people to find these activities rewarding, it is provided anyway. Later, when the pecuniary incentives are suddenly withdrawn, the subjects lose all interest in the activity. Therefore, the introduction of pecuniary incentives undermines self-motivated behaviour, and it suggests that instrumental motivation (money) does not mix well with intrinsic motives. Lutz argues that “this is probably as close as we can get to empirical evidence on the existence of two different types of ‘utility’” (Lutz, 1993, p. 152).¹⁷

After analysing some criticism of consequentialism, we can move to deontology. Of course, there are many approaches which are based on deontology but we are focused on Kant. His approach is a perfect example for two reasons. Firstly, he is perceived as the most prominent supporter of deontology. Secondly, he puts reason on the pedestal and his argumentation will be especially useful in the discussion on the capability approach. In general, deontology is an ethical system which regards morality as a duty that ought to be followed. It is about following universal norms that tell us how people should behave and what is right or wrong. Deontology is about principles which are independent of consequences. Moreover, deontology is predominately based on reason in opposition to utilitarianism which is mainly based on feelings (pleasure and pain). Reason provides moral rules which are indicated by a human will which makes freedom possible. In deontology, a moral problem is perceived to be a rational one and can be resolved by finding the right moral rule (Van Staveren, 2007).

Kant is the most prominent representative of deontological ethics because of categorical imperative. It is a rule that judges whether some action or maxim has moral status. Categorical imperative establishes duties that must be followed to act morally despite the cost to the agent. In general, most people know only one formulation of categorical imperative, whereas Kant distinguishes three of them, see (White 2004). The first one which is most know is “act only according to that maxim whereby you can at the same time will that it should become a universal law” (Kant, 1993, p. 30). This imperative is based on consistency. If everyone lies (adopts a maxim of lying), no one will believe anything that anyone said. The second version is the Formula of Respect (Dignity of Persons): “act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means” (Kant,

¹⁷ For a survey of such research, see the citations in (Schwartz, 1990). See also (Frey, & Jegen, 2001; Frey, & Oberholzer-Gee, 1997; Gneezy, & Rustichini, 2000; Lepper, & Greene, 1978).

1993, p. 36). This form is founded on respect for the dignity of humanity. When people lie, they are using their humanity, as well as people whom they deceive, as merely a means to an end. This rule is not only a negative principle that prohibits using others as means but it also teaches one to take the other people's ends as one's own. The third version is the Formula of Legislation for a Moral Community. It is very similar to the first version of categorical imperative "so act as if your maxims were to serve at the same time as a universal law (for all rational beings)" (Kant, 1993, p. 43). The last version as the first underlines the importance of universalization. However, it emphasizes the legislative nature of the will. The third formulation leads us towards the 'kingdom of ends,' an ideal state of the world where every person's individual ends exist side by side in a harmonious, moral community.

In the common opinion, Kant is identified with strict duties that are connected with categorical imperative. However, Kant's ethical system is much more sophisticated and deals with happiness and other ends as well. To understand the richness of Kant's approach we need to analyse the Kant's distinction between two types of duties, perfect and imperfect (White, 2004). A perfect duty is one that must be followed, or one "which permits no exception in the interest of inclination" (Kant, 1993, p. 30). Such duties are mostly negative duties. For example, duties not to steal or not to lie. Perfect duties are mostly derived from the Formula of Autonomy or of Universal Law and they are founded on the logical inconsistency of the forbidden act. On the other hand, imperfect duties are those that people have some latitude in executing - "the law cannot specify precisely in what way one is to act and how much one is to do" (Kant, 1996, p. 152). Imperfect duties grant latitude in following them because they only prescribe ends that must be followed, like duties of self-improvement or beneficence. These are usually positive duties which assign general attitudes to be taken up, not specific acts to be performed. Happiness was treated by Kant as imperfect duty "to secure one's own happiness is a duty (at least indirectly); for discontent with one's condition under many pressing cares and amid unsatisfied wants might easily become a great temptation to transgress one's duties" (Kant, 1993, p. 12).

To conclude, perfect duties are perceived as duties of action, whereas imperfect duties are perceived as duties of ends. When perfect duties bound people to a certain action, imperfect duties demands the adoptions of ends that must be pursued. However, it does not mean exclusions of the other ends, even those connected with self-interest (White, 2004).

White (2004, 2006, 2011) presents an interesting attempt to incorporate Kant's duty-based ethics into the standard economic model of rational choice. He incorporated Kant's categorical imperative into the neoclassical model of choice by making use of Kant's distinction between perfect and imperfect duties. These two types of duties can be integrated into the model of constrained preference

satisfaction, because perfect duties are included as constraints (alongside the standard budgetary and time constraints), whereas imperfect duties are included alongside the agent's other options in the preference ranking. It is essential for White to perceive the concept of preference in the formal sense in which no hedonic desire is implicated by preferences in general. However, "some of them are based on self and other-interested inclination and the others are based on imperfect duty" (White, 2006, p. 7). Because the term preference is very loaded, White uses more neutral terms like "ranking" or "ordering" (White 2006, p. 7). In White's model when people maximize their utility (understood as an index of preference satisfaction) including above-mentioned premises, they are respecting their perfect duties (by observing the constraints) as well as following their imperfect duties. White (2006, p. 7) concludes that "as long as the duties are followed for the sake of duty and not for another reason, the agent is choosing her actions according to Kant's basic system of ethics".

Why it is so important for White to move from loaded term preferences to neutral term rankings? It is a case because many economists assume that people must be self-interested and with this assumption, a person cannot be moral in the Kantian sense. Moral action is solely motivated by duty. Therefore, people cannot be motivated by the argument that being truly honest will benefit them in the long run. Even if one gives to charity because one derives satisfaction from being moral (it feels good), it is not a moral action in a Kantian sense because acting based on desire, even if admirable, is not moral. To be moral, one's actions must be performed out of duty, not merely according to duty. To sum up, we cannot perceive perfect duties as preferences because perfect duties cannot be perceived within the maximization framework.

This distinction between perfect and imperfect duties is essential from White's perspective. He treats perfect duties as some kind of commitment (White, 2006, p. 3) that we stick to, no matter the consequences. Here, we find some similarity of White's perception of duty to intentions and precommitment strategies. All of them distinguish values and goals from utility which is identified with preferences. White uses this distinction because he wants to propose, within neoclassical economics framework, the alternative ethical system to utilitarianism. On the other hand, the distinction between values and utility is essential in the context of this book because it undermines the revealed preference theory and starts a debate about human happiness. However, both approaches are the same at one point. Their aim is not to eradicate the framework of preference ranking but to make a distinction between values and utility which are perceived as the same in value-free economics.

Many economists observe that value-free economics has problems with explaining ethically-motivated behaviors which are prevalent among people. Value-free economics tried to resolve that by introducing a "taste for morality"

(White, 2006, p. 1). It is argued that people get utility from helping others. However, this explanation fails to seize what is special about moral motivation and this argument was presented in detail in Chapter 2 in the context of economic approach. The models based on duty explain behaviours like voting or tipping which baffled economists who try to explain them by broadening the understanding of self-interest. White (2006) indicates that commitment (moral rules) can override considerations of self-interest. Over the last four decades, many scientists have been trying to model ethically-motivated behavior and commitment, see (Brekke, Kverndokk, & Nyborg, 2003; Dowell, Goldfarb, & Griffith, 1998; Etzioni, 1987, 1988; Goldfarb, & Griffith, 1991a, 1991b; Minkler, 1999; Rabin, 1995; Sen, 1977, 1987; Van Staveren, 2001). For an overview see Hausman et al. (2017).

In the next chapter, we will analyse one of the deontological approaches – the capability approach which tries to establish objective goals without limiting freedom. We will see whether we can have a cake and eat a cake.

Chapter 6

The capability approach

*Rationality is not only about maximizing;
it is also about reflecting on what to maximize*
(Hédoin, 2013, p. 19)

6.1. Introduction

A few words of introduction are necessary before analysing the capability approach. The capability approach is a flexible framework which can be interpreted in different ways. It is not a flaw because it was done purposefully by Sen who is a founder of this approach. However, it means that it is necessary to indicate for what purpose we use this approach. In this book, the capability approach is perceived as a deontological approach. Firstly, it treats freedom, which is connected with reason, as an intrinsic value. Secondly, Sen is a fierce critic of the consequentialist approach taken by economics. He underlines that choice alone is not enough to understand human behaviour. Some argue that the capability approach is mostly compatible with virtue ethics (Van Staveren, 2007). However, this perception fits better to Nussbaum (2011) who writes about human flourishing and is strongly influenced by Aristotle. In my book, the main focus is on Sen who is also influenced by Aristotle, but his approach is closer to deontology. It is especially noticeable in the context of neoclassical economics which Sen criticizes. This is the point of departure for the capability approach. The purpose of this chapter is to use the capability approach as a tool to give the primacy to well-being and freedom. However, both of them are much more sophisticated than similar concepts used in neoclassical economics.

6.2. The criticism of neoclassical economics

Before introducing the capability approach, Sen had been comprehensively criticizing neoclassical economics for years. He especially criticized the revealed preference theory and neoclassical economics' understanding of rationality. Some of this criticism was presented in Chapter 3. However, it is vital to

shortly summarize the previous points and have them in one place because the capability approach is built on this criticism. Sen (1973) argues that there is a deep conceptual ambiguity at the core of the economic theory of rational choice. Neoclassical economists use the concept of preference to accomplish three different tasks: (1) to describe a person's choices, (2) to represent whatever motives underlie a person's choices, (3) to represent a person's welfare. For Sen, one concept of preferences, which is seen at the moment of choice, could not do so many completely distinct things. Therefore, he replaced one concept of preferences with three distinct concepts (choice, underlying motive, and welfare). Secondly, he proposed an alternative model of human behaviour – commitment (Anderson, 2001).

The second large conception that Sen criticizes is economic rationality which is a fundamental assumption for revealed preference theory. Sen (2002, p. 19) indicates that there are three different perceptions of the rationality in economics: (a) rationality as internal consistency of choice, (b) rationality as self-interest maximization, (c) rationality as maximization in general. Sen criticizes each of these assumptions and this criticism was already presented in this book. The first assumption was criticized in the context of menu-dependence. The second assumption was criticized in the case of commitment. The third assumption refers to instrumental rationality. It was demonstrated to what absurd conclusions not reflecting about goals leads.

For Sen the assumptions behind revealed preference theory and rationality have implication in real life. In welfare economics, the concept of preferences is used to (1) make judgments about individual welfare, (2) make judgments about the overall good of society, (3) articulate a principle of rational choice (utility maximization). Sen argues that none of the three concepts of preference (choice, underlying motive, perceived self-interest) measures well-being accurately. Moreover, preference identified with choice (revealed preferences) fails to do so as well (Anderson, 2001, p. 3). For Sen the main reason for that is the fact that people have motives that are wider and even counter to their self-interest. He called these motives commitments. In Chapter 2, the terms 'sympathy' and 'commitment' were introduced. However, they must be analysed further because the concepts of sympathy and commitment are essential for Sen who, as was indicated earlier in this book, characterizes them as:

The former corresponds to the case in which the concern for others directly affects one's own welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment. (Sen, 1977, p. 326)

The existence of commitment is essential because it demonstrates that we are social animals and we care for the other people and this fact is not fully taken into

consideration by value-free economics. Sen makes this point explicit when he writes: “I would argue that the philosophy of the revealed preference approach essentially underestimates the fact that man is a social animal and his choices are not rigidly bound to his own preferences only” (Sen, 1973, p. 253). Moreover, the concept of ‘commitment’ demonstrates that people do not always try to maximize their well-being. Therefore, individuals’ choices do not always reflect their personal welfare. It is not uncommon that people can distinguish between the questions: “Which is better, all things considered?” and “Which is better for me?” (Hausman, 2011, p. 7). Individuals’ ability to perceive the situation or their interest from the perspective of ‘impartial spectator’¹ is essential for Sen. It means that reason can be objective in contrast to emotions.

In general, Sen criticizes the standard notion of economic welfare because it is based only on utility and it does not contain other values. He writes: “Welfarism is an exacting demand, ruling out essential use of any non-utility information” (Sen, 1979, p. 478). However, his criticism refers not only to welfarism but to happiness as well. As Sen (1985c, pp. 188, 189) indicates:

Happiness has two basic problems in its claim to stand for well-being. First, as it is interpreted in the utilitarian tradition, happiness is basically a mental state, and it ignores other aspects of a person’s well-being. If a starving wreck, ravished by famine, buffeted by disease, is made happy through some mental conditioning (say, via the ‘opium’ of religion), the person will be seen as doing well on this mental-state perspective, but that would be quite scandalous. Second, as a mental-state concept, the perspective of happiness may give a very limited view of other mental activities. There are mental states other than being just happy, such as stimulation, excitement, etc., which are of direct relevance to a person’s well-being. Furthermore, mental activities involve valuation of one’s life – a reflective exercise – and the role of valuation in the identification of the person’s well-being obviously cannot be seen in terms merely of the happiness that such reflection itself creates. It is hard to avoid the conclusion that although happiness is of obvious and direct relevance to wellbeing, it is inadequate as a representation of well-being.

Sen highlights the two main problems of the hedonistic approaches towards happiness (e.g., subjective well-being approach). They are only reminded here because they were presented in Chapter 3 in more detail. Firstly, people’s preferences are malleable.² Secondly, happiness approach seems to ignore the importance of freedom as an intrinsic value.

The criticism towards welfarism makes Sen one of the main characters in the debate on human development, and this is probably the area where the

¹ Introduced by Hume and Smith. More about this conception in (Ostapiuk, 2017b).

² This critique was earlier presented by Berlin. He argues that when people’s lives are characterized by the obstacles and deprivation, they avoid important desires by retreating to their “inner citadel” (into themselves), because this give people some psychological security (Berlin, 2002, pp. 181–187).

capability approach has the most significant impact. Both human development and capability approach indicate that we should look more holistically at people's well-being and that economic growth should not be the only goal of development. Sen's approach is connected with the HDI index and with the attempts to measure human well-being by the other measures different than GDP (Godłów-Legiędź, 2008; Stiglitz, Sen, & Fitoussi, 2010). Sen is also well-known for using the capability approach to reduce poverty worldwide.

The capability approach is also prominent among scholars who deal with justice and the question how to distribute goods. For example, Sen (1999) uses the capability approach to determine what we should do in the following situation. The woman is hiring an unemployed labourer to work in her garden. There are three applicants, each of them would do much the same amount of work for the same fee. The employer is a reflective person and she wonders whom of these people should she hire. Sen imagines that the woman asks herself how by choosing from the applicants she can do the most good. Should she choose the poorest candidate (reduce poverty)? Should she choose the candidate who would gain the most happiness from being employed (increase happiness)? Alternatively, should she choose the candidate for whom the job would make the most significant difference to "the quality of life and freedom from illness"? (Sen, 1999, p. 54). The story is about alternative ways of distributing a resource among three individuals. The first approach is income-egalitarian, the second is about classical utilitarianism, and the third one is about capability approach. Even though in this section the focus is put on individuals and their well-being, not distribution, the understanding of capability approach also provides an answer to the issue of justice.

6.3. Functionings and capabilities

It can be useful to give a definition of the capability approach first and later explain what does it mean. So, the capability approach is a theoretical framework which is based on two normative assumptions. First, freedom to achieve well-being is of primary moral importance. Second, freedom to achieve well-being is to be understood in terms of people's capabilities, that is, their real opportunities to do and be what they have reason to value (Robeyns, 2016). To understand what each of these statements means we need to introduce the terminology used by Sen ('functionings' and 'capabilities'). At first, they seem unintuitive but one needs to comprehend what they mean to understand the capability approach.

Sen uses the example of a bicycle to explain the idea of functioning. A bike can be used as a means of transport or it can be sold. One can run alongside it while someone else is pedalling, etc. Therefore, a bike is a physical mean of

achieving something. It can be the ability to get around but also going out with friends, or getting into training for a cycle race, etc. The functioning is whatever each person achieves by using a bicycle (Ballet et al., 2013, p. 11). Functionings can be divided into ‘being and doings’ which are the various states of human being and activities that people can undertake. For example, beings are: being educated, being illiterate, being well-nourished, being undernourished, being part of a circle of friends, being part of a gang. The second group ‘doings’ are, for example, caring for an older parent, taking drugs, travelling, eating animals, playing with children, arguing.

These examples should indicate that many functionings can be described either as ‘being’ or as ‘doing’. We can state that a person is housed in a pleasantly warm house, or that this person does consume lots of energy to keep her house warm (Breton, & Sherlaw, 2011). On the other hand, many functionings are much more straightforward and they are either being or doing, for example, being healthy or eating animals. Secondly, ‘functionings’ is a theoretical category and in itself it is morally neutral. Functionings can be unambiguously good (e.g., being in good health) or bad (e.g., being injured). Some functionings are stable. Being well-nourished demands more or less the same amounts of commodities (food and health services). However, the other functionings depend on the society in which we live. One of Sen’s favorite examples of such functioning is ‘appearing in public without shame’ which is an example given by Smith. He observed that working men in his times wore linen shirt to work because without it they felt ashamed, whereas working men in ancient Egypt or Rome did not have shirts and it did not cause psychological distress (Sen, 1992, p. 115, 1999, pp. 73, 74). This example shows that some functionings differ in every society. This relativism means that the capability approach is a very flexible framework. Functionings not only differ among different societies but also within one society. We can assess differently goodness or badness of some functionings because of the different values that we cherish. Robeyns (2016) to point these differences out asks:

is the care work of a mother who is caring full-time for her child a valuable functioning or not? A conservative-communitarian normative theory will most likely mark this as a valuable functioning, whereas a feminist-liberal theory will only do so if the care work is the result of an autonomous choice made against a background of equal opportunities and fair support for those who have duties to care for dependents.

For Sen this diversity of functionings is a positive feature. He argues that functionings which are “the various things a person may value doing or being” (Sen, 1999, p. 75) give us the possibility to assess social welfare better than in the case of utility because they take into considerations the differences in individuals’ values.

The next essential term in the capability approach is ‘capability’. It is a potential functioning, which implies that it is a potential achievement. The idea refers to ‘what one could do or be’ and it differentiates capability from functioning where ‘what one does and is’ (Breton, & Sherlaw, 2011, p. 151). For example, a functioning such as ‘to have friends’ would be translated into a capability “to be able to make friends if one wishes to” (Breton, & Sherlaw, 2011, p. 151). Sometimes capabilities are realized and become functionings. However, in general, capabilities are perceived as the set of alternatives that a person has. Moreover, these alternative options are real opportunities for being and doing different things (Foster, & Sen 1997, pp. 199, 200). Therefore, if one starves because of one’s religion (fast) and this is one’s choice, then we mean capability. However, if one starves, because one does not have money to buy food, it is not capability, because one does not have the opportunity to do otherwise. It was not one’s choice to starve. For Sen freedom is an essential part, because he identifies it with capability. When functionings are the part that constitutes well-being (achievements), capabilities are the freedom to choose and to accomplish these functionings. Capability as freedom is an end in itself. We should not treat capabilities as solely means to achieve functionings. Capability is also a reflection on functionings that are worth achieving and that can be achieved. If functionings are valuable for us and they constitute well-being, then capability represents a person’s freedom to achieve well-being (Sen, 1992, p. 49). To properly understand what capability means, we need to analyze the concept of freedom which is very specific and complicated in Sen’s approach. It is done in the next subchapter. Then, the term capability should be more approachable.

Before moving to the investigation of the concept of freedom, it is necessary to underline that the capability approach is focused on end not means as it is done in value-free economics. From this perspective, the capability approach is in line with Weber’s approach to values in science because all normative values are explicitly ‘put on the table’ and it is known what our goals of pursuit are. As in Plato’s *Meno* “if you don’t have any idea what you’re looking for, you won’t ever find it” (Nussbaum, 2011, p. 15). The capability approach deliberately relies on the means-ends distinction. In this approach, we always should be explicit when valuing something, whether we value it as a means to a valuable end or an end in itself. In the capability approach, the ultimate ends are people’s capabilities. Therefore, it asks if people are able to be healthy (capability) and if they have access to means (resources) for this capability such as clean water, access to doctors or protection from diseases (Robeyns, 2016). The primary reason why in capability approach we assume that it is better to focus on the ends rather than means is that people differ in their ability to convert means into valuable opportunities (capabilities) or outcomes (functionings) (Sen, 1992, pp. 26–28, 36–38). Because ends are what ultimately matter if well-being

is considered, the means only work as a good approximation of people's opportunities to achieve those ends on the condition that all people have the same capacities to convert these means into equal capability sets (Robeyns, 2016). To understand what does it mean exactly, we can get back and use the example of a bike. As was indicated before, we are not interested in a bike itself, but in what we can use it for. A bike enables the functioning of mobility – to be able to move oneself freely and faster than walking. Because people have different abilities, they will get more or less from the bike. The relation between a good (bike) and the achievement of certain beings and doings (functioning) is called 'conversion factor' (Sen, 1992, p. 100). It is a degree to which a person can transform a resource into a functioning. An athletic person who was taught to ride a bicycle when she was a child has a high conversion factor enabling her to turn the bicycle into the ability to ride aptly. Whereas a person with a physical impairment or someone who was never taught to ride a bike has a very low conversion factor. Therefore, the conversion factor represents how much functioning one can get out of a good or service; in this example, how much mobility the person can get out of a bicycle (Robeyns, 2016).

Being focused on ends has two additional advantages besides capturing inter-individual differences. Firstly, the valuation of means will keep the status of an instrumental valuation and will not change itself into the intrinsic valuation. For instance, money or economic growth will not be valued for their own sake, but only in so far as they contribute to an extension of people's capabilities. Secondly, by starting from ends, we do not assume that there is only one important means to that ends (as income) but we ponder which types of means are essential for the fostering of a particular capability (Robeyns, 2016).

6.4. Freedom as an intrinsic value

To begin with, it must be underlined that Sen treats freedom as an intrinsic value. In general, the capability approach is agnostic when it comes to a valuation with one exception, freedom. Sen criticizes Friedman's idea 'free to choose' (Friedman M., & Friedman R., 1980) because it values freedom for the utilitarian reasons (more freedom = more utility). This should make him an ally with libertarians who also perceive freedom as an intrinsic value which is independent of consequences. However, Sen differs from libertarians in the way he understands freedom. Libertarians, as well as most economists, perceived freedom in negative terms. In general, "negative view of freedom focuses precisely on the absence of a class of restraints that one person may exercise over another, or indeed the state may exercise over individuals" (Sen, 1988, p. 272). It is often called 'freedom from.' We can also perceive freedom in positive terms where we are concentrated on what people can achieve or what kind of life they can have.

It is sometimes called ‘freedom for’ (happiness, wealth, etc.). This distinction was famously made by Berlin (1958).³ Sen refers to this distinction:

if a person happens to be poor and hungry because of low real wages or unemployment, without his having been prevented (by the state or by some strong-armed individual) from seeking a higher wage or finding employment, then the person’s negative freedom may not have been, violated, even though his positive freedom from hunger is clearly compromised by circumstances. (Sen, 1988, pp. 272, 273)

Sen thinks that negative freedom used by economists is not enough and we should perceive freedom in positive terms. However, the distinction between positive and negative freedom does not represent accurately Sen’s view on this matter, see (Qizilbash, 2011). His understanding of freedom is much more sophisticated and it involves both positive and negative freedom. Sen (2002, p. 510) argues that there are three main features of freedom which are interconnected: (1) opportunity to achieve; (2) autonomy of decision; (3) immunity from encroachment. The first two are connected with positive freedom but to achieve them we need to have guaranteed the third one which is negative freedom.

As was indicated in the context of the capability approach, freedom concerns “the real opportunity that we have to accomplish what we value” (Sen, 1992, p. 31, 1999, p. 74). It does not include freedoms or opportunities that people can hold theoretically but they lie beyond their reach in reality⁴ (e.g., be as wealthy as Bill Gates). This is why in the capability approach more choices that people have are not always better. For example, “capabilities are, by definition, limited to functionings of value; they exclude evil or harmful functionings” (Alkire, 2005, p. 121). Sen also argues that having more options does not automatically lead to an increase in freedom. Firstly, a new option may not be one that we value. Secondly, even if the option is valuable for us, more options (freedom of choice) may crowd out our ability to live “a peaceful and unbothered life” (Sen, 1992, p. 63). Sen highlights that “indeed sometimes more freedom of choice can bemuse and befuddle, and make one’s life more wretched” (Sen, 1992, p. 59).⁵ In general, Sen is not against having more choices, but he thinks that freedom of choice does not always increase one’s freedom.

The second point, autonomy of decision, is also complicated in Sen’s approach. It refers to the distinction between freedom and control (Sen, 1982). Freedom includes “a person’s ability to get systematically what he would choose no matter

³ The idea of distinguishing between a negative and a positive sense of the term ‘liberty’ goes back at least to Kant, see (Carter, 2018).

⁴ Rousseau has similar view “truly free man wants only what he can do and does what he pleases” (Rousseau, 1979, p. 84).

⁵ Schwartz (2004) demonstrates that having more choices can have negative implications for people.

who actually controls the levers of operation" (Sen, 1992, p. 65). For example, if we had a choice, we would choose to work in a smoke-free environment. This is why a public programme to prohibit smoking in common work areas does boost our freedom. It would be a case even if we were not asked about this issue directly. In the absence of this public programme, we would not have the freedom to work in a smoke-free environment. Our freedom is enhanced even though the number of alternatives from which we can choose does not increase (it even decreases because we lose the freedom to smoke). This example shows that for Sen the direct control is not the only expression of freedom (Alkire, 2005, pp. 120, 121).

Although Sen argues that direct control is not necessary, he thinks that individuals' choices must be respected and they cannot be overridden by somebody else. For him freedom is not connected with decisive choice but with decisive preference. Sen (1983, p. 19) presents the example of a person whose preferences related to medical treatment are respected even if she is unconscious to show why preferences are more important than choice in the case of freedom. In this example, a person's preference for averting a certain treatment guides the doctor even if the doctor thinks that it will reduce the probability of the patient's recovery. Sen believes that a person's liberty is preserved in this case. This is why he is in favour of a broader conception of freedom in which the preferences, not the direct choices are vital. He makes this point explicit when he writes that "To see liberty exclusively in terms of who is exercising control is inadequate" (Sen, 1983, p. 19).

If we look at the example with the doctor and smoke-free environment, Sen's idea can look paradoxical. In the first case, the state can decrease the number of choices, whereas the doctor cannot change patients' choice even if it is considered harmful. However, there is no paradox here, because in both cases individuals' preferences are what really matters, and what people really want. In the example of a smoke-free environment, we want all along not to smoke, and in the case with the doctor (before falling into a comma) the patient does not want some treatment. In both cases, individuals' real preferences are respected. If we look at the example with smoke-free environment, we can notice that freedom is connected with well-being. However, it does not mean that Sen equalizes these values. Even though he thinks that well-being is fundamental, he argues that freedom is more important.

The example with the smoke-free environment should only demonstrate that not all choices increase individuals' freedom. However, it does not mean that the state should decide what is good for people irrespectively of what they want. Even if the action of the state improves people's well-being and provides a good functioning, in such case people will not enjoy the capability of functioning, only "good functioning fortune" (Pettit, 2001, p. 9). As was indicated before, having

an option of achieving functioning is distinctive from the capability of attaining those functionings. For Sen, the latter is the main criterion of evaluation “quality of life is to be assessed in terms of the capability to achieve valuable functionings” (Sen, 1993b, p. 31). Capability is so important because it is connected with freedom and in some respect guarantees it. Sen writes that “the freedom to lead different types of life is reflected in the person’s capability set” (Sen, 1993b, p. 33). Sen understands that well-being is important but freedom is important as well. Moreover, he thinks that some forms of well-being will be better for being freely achieved (Sen, 1993b, pp. 39, 43).

Why does Sen think that freedom has an intrinsic value and well-being does not? To answer this question, we need to look back at commitment and sympathy. The distinction made by Sen demonstrates that people have different motivations than their well-being. Therefore, freedom does not necessarily mean improving one’s well-being. Because a person’s goal can differ from a person’s well-being, “it is important to distinguish between a person’s pursuit of individual well-being and a person’s pursuit of overall agency goals, where this latter concept includes all the other reasons that motivate individuals to act in the ways they do” (Davis, 2003, p. 153). In this context, the agency is not being perceived as in neoclassical economics but, as Sen points out, in its “older and ‘grandier’ sense as someone who acts and brings about change, and whose achievements can be judged in terms of her own values and objectives” (Sen, 1999, p. 19). Secondly, if being free to seek different goals is valuable to people, it is essential to make a distinction between the freedom to achieve one’s goals and the freedom to pursue one’s goals (Davis, 2003, p. 153). These two distinctions result in four concepts of individual advantage which are indicated by Sen: (1) well-being achievement, (2) agency achievement, (3) well-being freedom, (4) agency freedom (Sen, 1985a). The first concept refers to the concern of neoclassical economics in individuals’ interest in satisfying their own preferences. The second concept concerns individuals’ ability to achieve goals that do not need to involve their own well-being. The third concept concerns individuals having the freedom to pursue their own well-being. The fourth concept concerns individuals having the freedom to pursue all their goals, no matter if they are successful in achieving them (Davis, 2003, p. 153).

Although these conceptions are distinct from each other, sometimes they overlap. Sen argues that we cannot choose and focus on one of these conceptions and ignore the rest. Sometimes these different goals can be in conflict. For example:

if your riverside picnic is interrupted by the chance to rescue someone from drowning, then your agency freedom (and hopefully achievement) increases, because you can save someone’s life; but your achieved well-being diminishes, as you emerge cold wet and hungry. (Alkire, 2005, p. 122)

From the perspective of the above-mentioned examples it should be clear that Sen considers freedom as an intrinsic value. In his opinion, freedom gives people the possibility to have the life they have reason to value. The last part, 'reason to value',⁶ means that Sen underlines the importance of reason. He does not think that all individuals' preferences (desires) enhance freedom and well-being. He sees reason as the necessary condition for freedom and autonomy. Therefore, the last issue in our analysis of the capability approach is the question if and why reason guarantees freedom and autonomy. This question is especially important in the context of value-free economics in which rationality is perceived instrumentally and it can be identified with Humean approach.

6.5. Reason and autonomy

There are a few different perceptions of autonomy in philosophy. There is moral autonomy which started with Kant. It implies the capacity to ponder and to give oneself the moral law. There is political autonomy where one's decisions are respected within a political context. In this book, the focus is put on the third concept of autonomy – personal autonomy. In general, this idea refers to the capacity to be one's own person, to live life in line with motives and reasons that are perceived as one's own, not as the product of external forces (Dryden, 2018).

To discuss personal autonomy, one needs to assume that a true or authentic self exists. In fact, the idea of autonomy is strictly connected with a discussion about self and its constitution. As was pointed out in the discussion on personal identity, the idea of one self should not be taken for granted. Nowadays, many scholars support the socially constituted view of self. They argue that we cannot talk about autonomous 'true' self because we cannot consider agents in abstraction from social and relational commitments and contexts. For example, communitarians like Michael Sandel criticize the concept of autonomous self which is a fundamental assumption in liberal political theory. They argue that a person is embedded and shaped by social values and commitments. People do not choose their commitments and values from the position of already being autonomous individuals. The autonomous self does not exist prior to the commitments and values that constitute the basis for its decisions (Sandel, 1998). The idea of an autonomous self, which is based on reason, was also criticized by Romantics who reacted against the universality of reason introduced by the Enlightenment (e.g., Kant). They had their own conception of individuality in which passions are more important than reason in developing one's own unique self. This focus on emotions is also critical in phenomenological and existential philosophy where being 'authentic' is highly praised (Adorno, 1973;

⁶ He used it 29 times through the book (Sen, 1999).

De Beauvoir, 1949; Hegel, 2002; Heidegger, 1962; Kierkegaard, 1980, 1983, 1992; Sartre, 1992; Taylor, 1992).

These approaches which criticize personal autonomy based on reason will not be developed further for two reasons. Firstly, when the term 'emotion' is used in this book, it does not refer to some broad, metaphysical sense like inner life. It refers to the situations when people fall into temptations and smoke or eat unhealthy food, which are chosen against their better judgment (reason). Secondly, this book does not have any pretension to propose a new metaphysical theory of self. The goal is much humbler. It is to provide a distinction between reason and emotions which is not underlined in value-free economics.

The discussion between emotions and reason is very old and it started with philosophy itself.⁷ However, we deal with only two great philosophers, Kant and Hume, who are nevertheless essential in the context of this book. The first one is identified with reason and can be perceived as a supporter of value-based and goal-oriented economics in which we discuss human goals and values. On the other hand, Hume makes a case for emotions and can be perceived as the father of neoclassical economics which is value-free, because it is not entangled in the discussion about human goals. Some of the problems with Humean approach were indicated in a discussion on weakness of will. Now the question concerns even bigger issues: personal autonomy and free will.

In the first chapter, Hume was presented as a champion of instrumental rationality. He writes that "[r]eason alone can never be a motive to any action of the will", and "reason is, and ought only to be the slave of the passions" (Hume, 1896, pp. 282, 283). For Hume, passions in themselves cannot be called irrational or unreasonable because:

a passion must be accompany'd with some false judgment, in order to its being unreasonable; and even then 'tis not the passion, properly speaking, which is unreasonable, but the judgment [...] a passion can never, in any sense, be call'd unreasonable, but when founded on a false supposition, or when it chuses means insufficient for the design'd end. (Hume, 1896, p. 284)

What Hume is saying is that for an act to be irrational, the desire which was a cause of the action must be founded on false belief. The only way for an act to be irrational is to be ineffective (not to use proper means to achieve an end). Therefore, we can only say that a drunk is irrational not because he is drunk but because he drinks apple juice instead of vodka. Effectiveness is one of the axioms of value-free economics and it is the only way for value-free economics to judge some action as irrational.

⁷ Plato's famous dichotomy between reason and emotions, and his allegory of charioteer and horses (Plato, 1999).

In the instrumental view of rationality, the action is finally determined by psychological states. Therefore, an instrumental theory of choice presents a casual explanation of human action which has the same structure as the explanation that can be found in natural sciences. Kant understands that we need to accept this approach to explain the choices made by people because they are part of the physical world and we need to perceive their actions as having physical explanations (Sugden, 1991). However, Kant thinks that when people reason they have to perceive themselves as autonomous. We must conceive ourselves as able to form beliefs and reach conclusions which are not determined by external causes. Therefore, when people reason what actions to take, they must perceive themselves as being able to determine their actions. Kant thinks that reason has to regard itself as the creator of its principles which are independent from the alien influences. From this perspective, Humean passions are perceived as alien influences, because it is impossible to reason coherently about what to do when we think that our choices are fully determined by our psychological states.⁸ However, the purpose of Kant is not to prove that people are really autonomous. He argues that we cannot engage in any form of reasoning without presupposing our own autonomy (Sugden, 1991, p. 6).

The second feature of Kantian autonomy is the fact that an autonomous individual is one whose actions are governed by the laws that one imposed on oneself. Kant agrees with Humean instrumental rationality and he perceives it as the source of hypothetical imperative (if you want to achieve X, then do Z). However, there is another form of rationality which refers to categorical imperatives (do X, regardless of your wants). From this perspective, a choice is rational, if it is prescribed by some principle which can be a universal law for all rational agents (Sugden, 1991). We can find the analogy here with White's model of Kantian ethics in neoclassical economics framework in which perfect duties are connected with categorical imperatives and imperfect duties with hypothetical imperatives (White, 2004). The conclusions from this model were the same as in the case of rationality. There is a qualitative difference between moral laws and preferences as well as between rationality and instrumental rationality. In the end, Kantian categorical imperative cannot be embraced by the traditional economic framework.

For Kant reason is strictly connected with freedom and autonomy. Why can emotions not be perceived in the same way? Can we even assess whether emotions are rational or irrational?⁹ A negative answer to this question comes from two premises. Firstly, emotions being predominately involuntary cannot be perceived as actions. Secondly, only freely chosen actions can be assessed to

⁸ This state is called 'heteronomy' by Kant (1993).

⁹ The discussion about rationality of emotions is very extensive and there is no possibility to present it in this chapter (see (De Sousa, 1990)).

be rational or irrational (Elster, 1996, p. 8). Each of these premises was questioned but the explanations are far from being convincing. The first premise (emotions are involuntary) was questioned by the authors who argue that emotions are chosen by people (e.g., Sartre, 2015; Schafer, 1976; Solomon, 1980)). For them, emotions are actions in the same way as mental calculations are actions. Elster gives a commonsensical objection to these theories. He writes: "If we can choose our emotions, why do we not choose to be happy all the time? And why would anyone ever choose to be sad?" (Elster, 1996, p. 8). The second premise seems to be at variance with the idea of rational belief. People form beliefs, but they do not choose them. Actually, it is commonly accepted that the idea of deciding to believe is self-contradictory (Williams, 1970).

Frankfurt (1971) also argues that emotions (desires) do not make people free. As was only briefly indicated in Chapter 2, the condition when people are tossed by their desires Frankfurt called "wanton" (Frankfurt, 1971, p. 11). The predominant feature of a wanton is that it does not care about people's will. The desires move people to do a particular thing without asking them either that they want to be moved by those desires or that they prefer to be moved by the other desires. For Frankfurt 'wanton' is the feature of young children and animals. We do not perceive wanton as having freedom of will even though we recognize that animals can be free to run in whatever direction they want. Therefore, Frankfurt (1971, p. 9) argues that "having the freedom to do what one wants to do is not a sufficient condition of having a free will". Here, Frankfurt makes a distinction between being free to do what one wants and being free to decide what one wants. These concepts are completely different:

A person who is free to do what he wants to do may yet not be in a position to have the will he wants. Suppose, however, that he enjoys both freedom of action and freedom of the will. Then he is not only free to do what he wants to do; he is also free to want what he wants to want. (Frankfurt, 1971, p. 14)

In order to explain this difference, Frankfurt provides the idea of 'second-order desires' which was introduced in the second chapter briefly. For Frankfurt (1971), the capacity to have second-order desires (see meta-ranking (Sen, 1977)) distinguishes humans from other forms of life. Moreover, free will should be perceived as the ability to choose in accordance with one's second-order desires (called also 'metapreference,' 'metarankings,' or 'preferences over preferences'). It means the ability to construct the preferences that one would prefer to have.¹⁰

¹⁰ However, higher preferences do not automatically imply that they are better and people should follow them, Example of that is presented by Hollis (1983). Suppose a "conflicted" gay male who prefers men to women but prefers to prefer women because he is socially pressured. This person has a third-order preference for men. In such cases, Hollis argues, the agent would be better off trying to adjust his metapreference so as to bring it in line with his preference (as well as with his third-order preference)

Metapreferences and the idea of hierarchical self are frequently used by philosophers in debates on personal autonomy. When we look at Frankfurt's second-order desires, it is possible to see that this approach is much closer to Kant than Hume. Although individuals have preferences (first-order desire) as in Humean approach, they can reason and reflect on their preferences, and this is where the 'true' self lies. However, metapreferences approaches are not as strict as Kant's approach, because they allow the influence of society in shaping values (see (Sen, 1973); to know more about differences between Kant's and Sen's approaches see (Pauer-Studer, 2006)).

In the 1970s, Frankfurt with Dworkin were among philosophers who started the contemporary discussion on personal autonomy. The goal was to decide what kind of individual freedom should be protected and how it can be presented from the perspective of contemporary conceptions of free will. Their conclusion was that people's actions are worth protecting, if they are ingrained in values, general commitments and objectives that people have. On the other hand, they are not worth protecting, if they go against those values, objectives, and commitments. Both Frankfurt and Dworkin use hierarchical models in which an agent is autonomous only if her first-order desire is accepted by a second-order desire which is identified with volition (Frankfurt, 1988, pp. 12–25). This approach is neutral when it comes to the origins of higher-order desires. It does not exclude the desires and values that are socially constructed. The cause of such desires does not matter, only the agent's identification with them (Frankfurt, 1988, pp. 53, 54). Autonomy concerns our ability to examine and ask whether we do identify with our desires or whether we may wish to override them (Dworkin, 1988). 'We', in this case, is established by our higher-order preferences. Dworkin as Frankfurt speaks of them as the agent's "true self" (Dworkin, 1988, p. 59).

Identifying freedom with reason and reflection about desires is a common theme in philosophy. In general, freedom refers to the self-determination of a subject who possesses both reason and will (Ballet et al., 2013, p. 4). More fundamentally, freedom refers to the choice of a particular way of life that reflects one's values and commitments. It is a different account than negative freedom used in neoclassical economics where it is enough for people to act upon their autonomous preferences. Therefore, the concept of freedom presented in this book is the choice of values and commitments that reflect the life one is living (Ballet et al., 2013, p. 6). We need to keep in mind that self-determination does not mean the choice alone as it is in neoclassical economics (revealed preference) but the capacity to choose. In other words, the capacity to evaluate and re-evaluate one's choices (Ballet et al., 2013, p. 6) which can be identified with precommitment strategies and intentions. Sen's concept of self and personal autonomy is very similar to the above-mentioned philosophical concepts. For him the concept of self or person which is connected with commitment is explicitly reflexive.

This means that individuals are able to engage in reasoning and self-scrutiny. Sen (2002, p. 36) writes that:

A person is not only an entity that can enjoy one's own consumption, experience, and appreciate one's welfare, and have one's goals, but also an entity that can examine one's values and objectives and choose in the light of those values and objectives.

Moreover, Sen's perception of personal identity (self) which relies on reflexivity is essential for morality. He argues that only self-scrutinizing and reasoning individuals who are isolated from the instrumentality of self-goal choice can make commitments to others. It is so because only individuals who self-reflect are able to recognize the non-instrumental and intrinsic value of social rules (Davis, 2005, p. 23). These rules do not serve as means to some ends but can be ends in themselves.

Self-scrutiny is essential for Sen, because it can also constitute personal identity. It takes distinctness of people as the essential element in any account of personal identity. It is so because the act of taking oneself as a separate object of deliberation is equal to perceiving oneself as distinct and independent (Davis, 2005, p. 23).¹¹ Although Sen argues for the primacy of reason as it was done by Kant, he differs from him considerably, because the Noble Prize winner includes society in shaping values. For Sen there is not one strict path to establish values. People are social animals and their values are shaped by the society in which they live. Sen (2002, p. 215) writes that:

Community, nationality, race, sex, union membership, the fellowship of oligopolists, revolutionary solidarity, and so on, all provide identities that can be, depending on the context, crucial to our view of ourselves, and thus to the way we view our welfare, goals, or behavioral obligations.

At the end of the discussion on personal autonomy, we get back to the issue of personal identity. These two concepts are inextricably linked. In general, to indicate individual we must be able to initiate some course of action because this ability distinguishes a source from which something occurs. In neoclassical economics, agents act upon tastes that are imposed on them (they act upon the preferences they have), which makes it problematic to detect personal autonomy or identity in this approach.¹² Sen's idea of self is entirely different because he perceives agents as self-scrutinizing beings that can assess their desires by the overall values and goals and according to these values and goals. For Sen this is real freedom. Having in mind the discussion on personal identity and autonomy it seems justified to conclude that economics concepts

¹¹ We can see the influence of Smith on Sen once again. Self-securitization can be identified with Smith's 'impartial spectator' which also takes other people into consideration (see (Ostapiuk, 2017b)).

¹² The point that agents in neoclassical economics do not have freedom was made by Lawson (1997).

of personal identity and autonomy are too simplified and they do not reflect the intricacy of the long-lasting debates in philosophy.

6.6. The criticism of the capability approach

Over the last years, the capability approach has attracted a lot of attention from different fields of science. As it is not uncommon, with popularity comes criticism. In this section, some part of this criticism is presented. There are two reasons why not every criticism of the capability approach is examined. Firstly, much criticism refers not to capability approach itself but to the implications of this approach. As was indicated before, Sen deliberately designed a flexible framework which can be used in different ways. Therefore, the problems with his framework frequently concern the implementation of the capability approach and not the framework itself. Secondly, the criticism predominately concerns a social aspect of Sen's approach. The focus in this work is on individuals, so there is no need to broaden the scope here.

There are two main kinds of criticisms of the capability approach that will be analysed. The first one concerns a philosopher Martha Nussbaum. However, this criticism can be perceived more as the expansion of Sen's work than its dismissal. Nussbaum indicates three differences that distinguish her capabilities approach from Sen's approach. Firstly, she criticizes Sen for not dismissing cultural relativism explicitly. Although Sen has sympathy for universal norms, he was not able to reject cultural relativism completely (Nussbaum, 2000). Secondly, Nussbaum criticizes Sen for not establishing his theory in the Aristotelian framework in which human functioning is essential. Nussbaum argues that Sen uses the structure of Aristotelian approach but he did not make it explicit. If he did, he would be able to provide the lists of central capabilities (see (Kleist, 2018)). Nussbaum creates such a list herself, which is one of her most significant contribution to the capability approach. In general, she is mostly interested in a theory of basic social justice and in this area she used the capability approach. She used the list of central capabilities to develop a theory of 'universal fundamental political entitlements' where entitlements are given, by a list of ten central capabilities: (1) life, (2) bodily health, (3) bodily integrity, (4) senses, imagination and thought, (5) emotions, (6) practical reason, (7) affiliation, (8) other species, (9) play, and (10) political and material control over one's environment (Nussbaum, 2011, pp. 33, 34). From the beginning, the list has been fiercely debated, which was encouraged by Nussbaum who pointed out that the capabilities on the list can be changed.¹³ However, the list has not been changed since its beginning.

¹³ At the beginning, Nussbaum (1995) argued that her list was static but later she changed her mind.

In general, it must be stated that Nussbaum's criticism towards Sen is the result of the Aristotelian tradition from which she descends. Nussbaum is focused on the goals (*telos*) and human flourishing. This is why she tries to switch the main focus in Sen's capability approach (freedom) towards well-being. It is visible when she wants to dismiss the distinction made by Sen between well-being freedom and agency freedom. She argues that "the distinction is obscure and not useful to one who, like Sen, has rejected (on good grounds) utilitarian notions of well-being" (Nussbaum, 2011, p. 200). Nussbaum argues that this distinction is not needed, because in her approach we are focused on capabilities rather than functionings and the practical reason has a fundamental place in the list of capabilities. She writes that "because what is valued is the freedom to do or not to do, agency is woven throughout" (Nussbaum, 2011, p. 201).

However, many philosophers who deal with the capability approach, state that the distinction between well-being and agency is useful. For instance, agency may refer to particular sacrifices one might want to make of one's own well-being out of commitment to collective values (e.g., the environment) or out of commitment to the value of the quality of life of others (e.g., the decision of an adult child to care intensively under difficult circumstances for her terminally ill parent) (Robeyns, 2011). In these cases, people can easily distinguish between the course of action which is the best for them (well-being) and which is a thing they value the most (agency).

In general, the differences between Sen and Nussbaum are not so significant and we can perceive these approaches as complementary rather than opposing each other. Both Sen and Nussbaum have a high regard for each other. The differences in theories can be a result of the different goals that Sen and Nussbaum set for themselves. Sen is more interested in the assessment of the quality of life, whereas Nussbaum tries to implement her theories in social justice. In this book, Sen's approach is used because it gives primacy to freedom. It guarantees pluralism where everyone can build their own system of values. Thanks to that, we should not fall into a dogmatism ('basic' value judgements (Sen, 1970)). Freedom is a good foundation for the different normative approaches and Nussbaum can be treated as one of the thinkers that implements Sen's framework.

The second kind of criticism towards Sen's approach comes from a completely different place. It concerns a threat to freedom because of the activity of the state. Especially these words written by Sen make libertarians very uncomfortable:

Being free to live the way one would like may be enormously helped by the choice of others, and it would be a mistake to think of achievements only in terms of active choice by oneself. A person's ability to achieve various valuable functionings may be greatly enhanced by public action and policy, and these expansions of capability are not unimportant for freedom for that reason. (Sen, 1993b, p. 44)

This paragraph causes many problems, because Sen has never really explained in detail how the democratic process and public reasoning could select and enhance valuable functionings. The process of choosing between functionings is important, because we know that some of them do not enhance individuals' well-being and freedom. Thus, they do not deserve social attention (e.g., smoking). The critics of the capability approach identify it with an Aristotelian approach in which the most crucial issue is human flourishing. From this perspective, it does not seem controversial to judge negatively capability to smoke because it does not enhance human flourishing (Breton, & Sherlaw, 2011). This kind of conclusions come from Sen's perception of capabilities in which people can "lead the kind of lives they value – and have reason to value" (Sen, 1999, p. 18). There is a potential ambiguity around the expression 'reason to value.' The critics of the capability approach used this ambiguity to question whether the capability approach, despite democratic deliberation advocated by Sen, could cause the infringement of personal freedoms like the freedom to smoke (see (Sugden, 2006, 2010)). Sugden makes this point explicit when he writes that "my unease about Sen's proposal is that it licences collective decisions that override some individuals' actual desires about how to live their own lives, in favour of other people's judgements about what people have reason to desire" (Sugden, 2006, p. 41). Sen (2006) dismissed such a critique. Qizilbash (2011) comprehensively explained why the Sugden's criticism is the result of the misrepresentation of Sen's approach.

There is no need to delve into these dense explanations. However, a few points must be made to address some of Sugden's concerns, because they are of great importance for public life in general. Firstly, it is true that the capability approach can be ambiguous, but it is a result of its flexibility. In the case of smoking, Sugden's concerns can be justified, because Sen argues that the state which creates a smoke-free environment does not diminish but boost freedom. It is possible because we do not value choices for their own sake. They must be in line with our 'decisive preferences' ((Sen, 2002, ch. 20, 21), see (Pettit, 2001)). The problem that is indicated by Sugden is how we can know what individuals 'really' want if we do not observe their choices. Also, how the state can decide which preferences are valuable and guided by reason? Although the unease of Sugden is justified, the conclusions that he draws (infringement of personal freedom) are not.

It is a case because freedom in the capability approach has an intrinsic value and it is highlighted by Sen. We can observe that in the case of the doctor who needs to respect values and commitments of the unconscious patients even though they are bad for these patients and decrease their well-being (Sen, 1983, p. 19). This example indicates that the capability approach is more about having appropriate information and knowledge to decide what people's values

are rather than to decide which values are desirable for them. This is why Sen underlines the importance of the education system which increases people's capabilities.¹⁴

The example of a comma is extreme, but one can imagine many situations in which people stick to the values that diminish their well-being. For example, imagine a conscious person who is Jehovah's Witness and she does not want the transfusion of blood because of the religious values even though it means death. There is a question whether values are the result of social pressure and they are connected with the malleability of preferences as it was in the case of Arabic women or they are real values which are important to this person. There is no clear answer to this dilemma. However, it seems plausible that Sen would accept the decisions of Jehovah Witness even if it meant her death because freedom has an intrinsic value. However, before that, he would probably try to rationally discuss this issue with the patient and make her well-prepared for a decision. Of course, there always will be the question how much knowledge is enough to be able to make a well-informed decision. However, as was indicated in Chapter 5, we cannot escape from relativism, and lack of definite answers is not a sufficient reason to reject some conceptions. From this perspective, a public discussion advocated by Sen can provide more capabilities because people will have more information to reflect upon. It goes without saying that establishing which conditions are proper for decision making is paternalistic. However, because people live together in a democracy where government exists these conditions must be established somehow.

The next problem with Sugden's criticism is that his libertarian position implies that we can only know what people really want if we look at their choices. However, one of the main goals of this book was to establish that people often do not choose the best options because of cognitive biases. Moreover, negative perception of freedom, as in libertarianism, is too simplified to capture the convoluted connection between freedom and well-being which the capability approach tries to underline. Libertarian position is very similar to neoclassical economics view with revealed preference theory. Both use the same assumptions about personal identity and they perceive freedom as 'free to choose.' Therefore, the criticisms of neoclassical economics presented in this book will apply to some parts of libertarian argumentation as well.

Although this criticism was presented, it concerns only individuals. The interaction between the state and individuals are much more complicated and Sugden's worries about infringement of freedom are justified. The discussion about happiness versus freedom is presented in the next chapter on the example

¹⁴ Being educated is described by Sen (1992) as a basic capability.

of libertarian paternalism. It is checked whether the state can guarantee both of these things at the same time. Moreover, the idea of libertarian paternalism is presented as the practical implementation of the capability approach.

6.7. Conclusions

Here I want to conclude the last three chapters. I do so, because they are interconnected and they form a part of a larger whole. The goal of Chapter 4 was to investigate the concept of multiple self. The implications of this concept suggest that both revealed preference theory and subjective well-being do not properly describe and explain human happiness. It is visible on the example of hyperbolic discounting where choices over time cannot be compared by utility measures on which two previous approaches are based. Moreover, hyperbolic discounting demonstrates that people have different selves with completely different preferences and goals. To justify the concept of multiple self, the philosophical discussion on personal identity was presented. The conclusion from this debate is that the conception of stable self over time which is used in neoclassical economics is dismissed by many philosophers who indicate the problems with this idea.

From this perspective and because of hyperbolic discounting it seems justified to distinguish between 'long-term human' and 'short-term human'. The former has long-term values and goals and it can be identified with reason and self-reflection. The latter is driven by emotions and temptations which are predominately connected with the short-term pleasures. Value-free economics assumes that the problem with different preferences over time does not exist because what people choose is what they want and it indicates who they are (revealed preference theory). However, comparing intertemporal choices by utility measurements seems impossible. This problem raises the question which self should be in control and which goals and values are more important. To answer this question, it is necessary to establish values that are independent of choice. The effective way to do that is to observe precommitment strategies in which the contest between 'long-term' and 'short-term' human is visible. The idea of precommitment is also connected with the discussion on weakness of will. The conclusion is that commitments and intentions are necessary to understand and accept the possibility of weakness of will. Without these conceptions many human behaviours do not make any sense.

Value-free economics is reluctant to accept the conception in which values are independent of choices because the revealed preference theory is very effective. It is one of the reasons why economics has a dominant position among the other social sciences (economic imperialism). In Chapter 5, the conclusions from the contemporary philosophy of science were presented to indicate that it is

impossible for scientists to establish the absolute Truth and economists are not the exception. Therefore, there is no methodological higher ground for positive/value-free economics and normative approaches are not intrinsically less scientific than positive approaches. It means that economics needs to investigate values. Weber is presented as a scientist who knew that scientists could not get rid of values in their research (Weber, 1949). This is why they need to rigorously scrutinize values and indicate them explicitly in their work. Economists need to put values on the table. Deontological approach, which shows values explicitly, was presented as the opponent of consequentialism which is used in value-free economics. Kant as a main figure of deontology was analysed to demonstrate that some parts of his theory (categorical imperatives) do not fit into the preference framework used in value-free economics. It means that morally or value-motivated actions cannot be fully incorporated into value-free economics.

In Chapter 6, the capability approach was presented as a way of establishing values independent of choice and that it can help us in better assessment of well-being. The capability approach criticizes conceptions of well-being proposed by both neoclassical economics (welfare) and economics of happiness (subjective well-being). However, the capability approach does not provide us with objective values besides freedom. Sen's approach is a flexible framework based on freedom which enables people to find their own personal values and goals. Although it is a very personal approach, Sen's statement that what matters is "our capability to lead the kind of lives we have reason to value" (Sen, 1999, p. 285) implies that our values are not random. Sen gives primacy to reason and self-reflection. This suggests that in some way he would support the 'long-term human,' not the 'short-term human' which is driven by emotions. Although Nussbaum extended Sen's approach by giving primacy to well-being, Sen's flexible framework is used in the book. The purpose of this chapter was not to indicate what is good for people as it was done by Nussbaum, but to show that economists need to choose normatively between selves. The primacy of reason indicated by Sen seems to support the author's conception of 'long-term human,' because reason guarantees both well-being and freedom.

In the next chapter, libertarian paternalism is presented to check whether it can provide both freedom and well-being and if the idea of 'long-term human' can be supported. Also, it is checked if it is even possible to discuss human goals and make decisions based on them in real life. Moreover, we will see whether the conclusions from previous chapters can be used to put libertarian paternalism on a more solid methodological ground. It can provide an answer to some criticisms of libertarian paternalism.

Chapter 7

Libertarian paternalism. The discussion on freedom and well-being

7.1. Introduction

In the previous chapters it was indicated, among other things, that we need to have a discussion about goals and what is good for people. Value-free economics was criticized from various perspectives. The revealed preference theory was the conception that was criticized the most. Mainly, it was pointed out that it is not always a good indicator of well-being. Value-free economics was also presented as a theory which was not descriptively accurate. Moreover, it is based on axiomatic assumptions which in the end lead the economic approach to be the degenerative programme. Value-free economics was also criticized from the epistemological perspective. It was argued that science should not be reduced to a process of building systems which are based on axioms and 'objective facts.'

This criticism concerns the theoretical foundations of value-free economics. In this chapter, we derive practical implications from the criticism and we introduce a normative approach that can resolve some of the above-mentioned problems. Underlining practical concerns is crucial, because the theoretical assumptions of value-free economics have a significant impact on the contemporary free-market society. Thanks to the mathematical models and unified methodology, economics has high predictive power which makes it commonly viewed as the most influential among social sciences. Economists are constantly invited to TV studios as experts. In general, they shape political and social reality in a greater degree than the other social scientists. The zenith of neoclassical economics' influence was in the 1980s when Regan and Thatcher implemented some theoretical ideas of neoclassical economics into reality.¹ After the 2008 crisis neoclassical economics experienced a backlash, but it has not changed the foundation of neoclassical economics (Fiedor, 2010; Ratajczak, 2017). Still, the assumptions of neoclassical economics about negative freedom and rationality (people know what is the best for them) are very influential. In most of capitalistic and democratic societies it is assumed that giving people choice is enough ('free to choose' (Friedman, M., & Friedman, R., 1980)), because people are rational and autonomous and they know what is the best for them.

¹ For historical analysis of this process, see (Mirowski, 2002).

In the end, value-free economics mainly leads to a normative political system – capitalism with the free market.²

It is important to underline that capitalism with its values has not been imposed on people. They really wanted it. Most people think that free will is about making choices which are unconstrained by the external factors. When one does as one wants (Monroe, & Malle, 2010). This view supports neoclassical economics' view of freedom which is perceived as negative freedom (freedom from). This is why many people argue that a free market provides freedom in contrast to the government. Also, Stillman found that people's understanding of free will is connected with morally responsible behaviour, achieving goals, self-control, deliberation, and conscious thought (Stillman, Baumeister, & Mele, 2011). It also means that people connect rationality with free will.

In a capitalistic system it is assumed that people do what is best for them, because they are rational. This assumption together with the perception of people as autonomous entities is the foundation on which contemporary liberal democracies are based on. The purpose of this book and this chapter is not to dismiss the assumptions of rationality and autonomy, but to criticize these assumptions to improve them.

The next reason for criticizing the consequences of value-free economics is the fact that the developed countries became much wealthier after the Second World War, but it did not make people happier (e.g., (Easterlin, 1974; Offer, 2006; Scitovsky, 1976)). The insight that consuming more goods does not guarantee happiness surprises only economists. The vast majority of philosophical approaches towards happiness argue that hedonism does not guarantee happiness. For them, it is a simplified version of well-being which contains much more than pleasures (Alexandrova, 2017; Bruni, & Porta, 2005, 2007; Bruni et al., 2008; Fletcher, 2015). Even hedonists like Epicurus or utilitarianists like Bentham understand that higher goods (e.g., intellectual pleasures) can give people more happiness rather than simple (sensorial) pleasures. Smith, the father of contemporary economics, also thought that the craving for material goods was not going to make people happy even if it was good for economy (see (Rasmussen, 2006)). The revealed preference theory was taken at face value in a capitalistic system. It is a problem because, as was indicated before, it was only proposed as a technical framework in order to make axiomatic assumptions of value-free economics work. It should never have been treated as a substantive theory of well-being (what is good for people), which is how it is predominately perceived in a capitalistic system.

² However, it is important to notice that it can lead to different system. For example, socialism proposed by Lange. On the other hand, it is possible to be against neoclassical economics and be in favour of the market as Hayek (thanks for this comment to Professor Godłów-Legiędź).

The technical assumptions of value-free economics seem to have a negative impact on society and people. Libertarian paternalism is presented as a normative conception that can help people with achieving both freedom and well-being. Besides the presentation of libertarian paternalism, one of the most important goals of this chapter is to provide a philosophical grounding to libertarian paternalism. It has not been sufficiently done by Thaler and Sunstein who were focused on the practical implications of their conception. The conclusions from the previous chapters are used to check whether some claims made by libertarian paternalism can be justified. Firstly, revealed preferences are not always a good indicator of well-being and this is why Thaler and Sunstein (2008) use the findings from behavioural economics and economics of happiness to discover what brings happiness to people. The conclusions from the discussions about precommitment strategies, weakness of will, and multiple selves are used as the arguments for libertarian paternalism. Moreover, libertarian paternalism is perceived from the perspective of the capability approach in which both freedom and well-being can be simultaneously satisfied. Furthermore, the criticism of libertarian paternalism is presented to indicate and establish the limits of this approach.

The next goal of this chapter is the discussion about freedom and happiness. The previous discussion on freedom and rationality is used in the context of libertarian paternalism which has a more sophisticated view about these issues than revealed preference theory. It is argued that sometimes we need to compare well-being with freedom and choose what is best for people. This is why freedom cannot be treated as a 'nonbasic' value (Sen, 1970) which cannot be discussed. In general, the discussion in this chapter is constrained to issues that were discussed previously in the book. The goal of this chapter is not an attempt to resolve the endless debate between libertarianism and utilitarianism.

7.2. Psychological background

At the very beginning of their book, Thaler and Sunstein make a distinction between 'econs' and 'humans' (Thaler, & Sunstein, 2008, p. 7). This distinction is the result of the criticism towards *homo economics* in its most idealistic form where it is assumed that people are fully rational and can compute all the information. Humans, on the other hand, are ordinary people who are frequently irrational and they are influenced by cognitive biases. This distinction is the result of the psychological knowledge that people use two different thinking systems. Thaler and Sunstein named their systems 'automatic' and 'reflective' which in psychology are frequently referred to as 'system 1' and 'system 2' (Thaler, & Sunstein, 2008, p. 19). Despite the differences these two systems are interconnected, which was indicated in Chapter 4. Table 2 shows the differences between these two systems.

Table 2. Comparing the automatic with a reflective system

| Automatic System (system 1) | Reflective System (system 2) |
|-----------------------------|------------------------------|
| Uncontrolled | Controlled |
| Effortless | Effortful |
| Associative | Deductive |
| Fast | Slow |
| Unconscious | Self-aware |
| Skilled | Rule-following |

Source: (Thaler, & Sunstein, 2008, p. 20).

‘System 1’ is mainly responsible for the cognitive biases and heuristics which are identified with irrationality. If we had used only ‘system 2’, *homo economicus* would not have been only a theoretical model. However, as was indicated earlier in this book, the constant deliberation about every prosaic choice would make people’s lives unbearable, because it would take too much time to make some decisions. At the beginning of the 1980s psychologists and behavioural economists were focused on the errors of the system 1 to show the unreality of *homo economicus* (e.g., (Kahneman, 2011; Thaler, 2015)). Despite this criticism, the conclusion is that people are inhabited by both *homo sapiens* and *homo economicus* and ‘system 1’ works together with ‘system 2’ (Kahneman, 2011).³

It goes without saying that nowadays most economists do not perceive people as hyper-rational agents (econ) and they know about bounded rationality and cognitive constraints. However, other social scientists think that most economists really believe in *homo economicus*. Over the last years, *homo economicus* started to be a strawman that many social scientists criticize (Chmielewski, 2011; Cohen, 2014; Kahneman, 2011; Sedláček, 2011). Even though nowadays very few economic scholars really believe in *homo economicus*, it is still a popular conception among the supporters of a capitalistic system. Thaler and Sunstein spend a lot of time to demonstrate that people are not perfectly rational. They use the findings from behavioural economics to make that point. This knowledge is well-known in the literature and was presented through this book. This is why it will not be expanded further in this section.

³ The necessity of system 1 was famously demonstrated by Damasio (2005) on the example of Phineas Gage.

7.3. Libertarian paternalism

7.3.1. Introduction

By the term ‘libertarian paternalism’ Thaler and Sunstein want to demonstrate that these two opposite systems can work together. The creators of libertarian paternalism have large ambitions and they perceived their theory as a ‘third way’ which can reconcile capitalism with socialism. For them *laissez-faire* and interventionism can work together. Thaler and Sunstein put their theoretical conception in a political context. They argue that the discussion between Democrats and Republicans which has been going from the time of New Deal (Roosevelt) does not make sense. From that time, we can observe a strong polarization both on political and social grounds. On the one hand, Democrats highlight market failures and they expect from the government to do something about them. On the other hand, Republicans underline the adverse outcomes of state intervention and they argue that a free market is ideal and it should not be constrained by the state under no circumstances. From this perspective, it seems that the consensus between these two camps is impossible. However, Thaler and Sunstein think that we can find common ground for these two opposite approaches and libertarian paternalism can be the ‘third way’ (Thaler, & Sunstein, 2008, pp. 252, 253).

Their theory contains both libertarian and paternalistic dimensions. Thaler and Sunstein argue that people should be free to do what they want. They use the famous wording presented by Friedman that people should be ‘free to choose’ (Friedman M., & Friedman R., 1980). The paternalistic dimension is the result of behavioural economics which indicates that people have cognitive constraints and they do not choose the best option. The important issue for Thaler and Sunstein is that people’s lives should be better. However, it does not mean that the state decides for people what is good for them. They should be better off by their own judgment (Thaler, & Sunstein, 2008, p. 5).

Libertarian paternalism is a weak, light and invasive type of paternalism. If people want to smoke cigarettes, eat too many sweets, choose the wrong health plan or not to save for retirement, then libertarian paternalists will not force them to do otherwise. On the other hand, this approach can be perceived as paternalistic, because ‘choice architects’ (later it will be explained what does it mean) try to push people in the direction of the choices that make people’s lives better by means of nudges. ‘Nudge’ is the title of Thaler’s and Sunstein’s book (2008). It is a vital part of the libertarian paternalism, because thanks to nudges choice architect can direct people towards the right decisions.

Libertarian paternalism can be explained on the example of parents who raise their children. The very term ‘paternalism’ descends from Latin word *pater*

(father) who can be identified with a strict parent who tells children what they should do and who does not give them the possibility to choose independently. On the other hand, parents can choose a stress-free upbringing where a child can do whatever it wants. Thanks to freedom it can learn to think and make choices independently. This approach should be identified with *laissez-faire*. The third option preferred in libertarian paternalism are parents that allow children to make their own choices, but simultaneously they establish the rules and norms to which children need to relate. Moreover, parents try to help children with taking the right decisions. Libertarians would question the adequacy of the above-mentioned comparison, because a grown person cannot be compared with a child. Parents can decide on behalf of children, because children have problems to make the right choices. Libertarians do not make a claim that everyone must decide for oneself. They give the example of children and psychotics for whom the decisions can be made by the others, because they cannot choose properly for themselves. However, for behavioural economics people have so many cognitive biases that they frequently do not choose in accordance with their own interest. This is why from behavioural perspective people sometimes can be treated as 'children' that should be helped. Of course, this approach results in the objections from libertarians that people are infantilized. This issue will be discussed later.

It is important to underline that paternalism has different forms and it is not necessarily as invasive as libertarians might think. Sunstein (2013b) shows the difference within the paternalistic position. One can imagine paternalism within which the influence is limited to the choice of means only. The second approach is paternalism within which the government decides about goals and which is much more invasive than the former. The example of 'means paternalism' is the government that encourages or even presses people to have the fridges that have low maintenance costs because saving money is exactly what people want (Sunstein, 2013b). More invasive version 'ends paternalism' is the government that banishes some sexual activities even though they are people who derive pleasure from these activities and want them (Sunstein, 2013b). Behavioural economics and the founders of libertarian paternalism are focused on 'means paternalism' not the latter. The next distinction is between 'hard' and 'soft' paternalism. 'Hard paternalism' contains methods like a prison, fine and other financial penalties. 'Soft paternalism' consists of warnings, default option and disclosure of information. In this context, nudges are the means of soft paternalism (Sunstein, 2013b).

7.3.2. Choice architecture

Thaler and Sunstein think that the state in libertarian paternalism should be a 'choice architect' who by using the default option directs individuals towards

decisions which would be taken if it were not for cognitive biases. The most frequently used example to understand how choice architect works is a person who is responsible for the arrangement of food in school cafeterias. The idea is to display at the beginning of a counter the healthy food and less healthy food closer to the cash desk. People want to eat healthily most of the time. However, they cannot resist the temptation and they reach for unhealthy food. Thanks to the proposed display of food, the temptation for unhealthy food is weaker, because it is harder to get them and individuals' strong will is not exposed to the temptation. In this case, people still can choose unhealthy food if they really want it. On the other hand, it is easier to stick to the commitments for people who want to eat healthily but fall into the temptation (Thaler, & Sunstein, 2008, p. 2). According to Thaler and Sunstein, this option of the arrangement of food is the best if one looks at the alternatives. They gave the example of Carolyn who is a director of food services for schools and she must decide how to arrange the food for kids:

1. Arrange the food to make the students best off, all things considered.
2. Choose the food order at random.
3. Try to arrange the food to get the kids to pick the same foods they would choose on their own.
4. Maximize the sales of the items from the suppliers that are willing to offer the largest bribes.
5. Maximize profits, period. (Thaler, & Sunstein, 2008, p. 2)

The first option is attractive but very paternalistic. The second option's impartiality can seem neutral but children in some schools will have a healthier diet than in other schools. The third option looks right, because we give children a choice. However, this system is not neutral, because the choice of children depends on the way in which the products are displayed. This is why it is so hard to know the true preferences of children. The fourth option is not interesting for us, because we are honest and we even did not think about cheating. The fifth option is interesting but problematic from the perspective of the person whose goal is the well-being of children (Thaler, & Sunstein, 2008, pp. 2, 3).

The essential thing about choice architecture is that there are situations in which something or someone always has an impact on a choice. This is why neutrality does not exist. In some situations, the state cannot remain neutral and paternalism is inevitable (Thaler, & Sunstein, 2008, p. 237). Therefore, it is better for the context of choice to be well-planned. This view challenges the libertarians' belief that it is the best to give people as many options as possible from which they can choose. Thaler and Sunstein demonstrate that more options can work in contrary to people's interests on the example of the privatization of the social security pillar in Sweden. People could create their own portfolios by selecting up to five funds from the 500 which were put on the list. The Swedes were

encouraged to make an active choice (the funds could advertise, every Sweden get a thick book with information about all funds). The system was introduced in 2000 and at the beginning two third of people chose the portfolios on their own, whereas one third did not choose them and they were assigned to the default fund chosen by the state. With time, it occurs that the profit from the default fund was forth times higher than the average profit from the independently chosen portfolio. This example shows that giving a free choice for people is not always the best option. Especially when a choice concerns some complicated issues that most people have limited or no idea about (Thaler, & Sunstein, 2008, pp. 145–156).

Default option is the method that can be used by ‘choice architect.’ It is one of the easiest, cheapest and most effective nudges by which the government can direct people. It means that when people have many choices, one of them is the default option chosen by the government. If people do not choose the other options, they are assigned to some default option. This option has a strong influence on people because of ‘*status quo* bias.’ This term was introduced by William Samuelson and Richard Zeckhauser (1988). People’s need to maintain *status quo* evinces in many situations. Most teachers know that students have the tendency to sit in the same places in the classroom even though it is not sanctioned anywhere. The consequences of *status quo* bias can be much more serious. The research conducted in the 1980s among professors who participated in retirement plans shows that more than half of them had never changed the way the payments were made. As a result, the decisions concerning a division of assets that were made many years before have never been changed. Even though the situation on stock exchanges and knowledge of future seniors has changed. *Status quo* bias can be easily extracted and many firms do it. The examples of that are free subscriptions to newspapers or the other long-term contracts in which at the beginning people get free products but later they need to pay for them. It turns out that many people subscribe and buy these products even though they do not use them. What works here is the inertia effect. People just do not want to change anything (Thaler, & Sunstein, 2008, pp. 24, 25).

Default option can have a major influence on people’s life, and organ donation is the best example of that. As it turns out, various countries differ entirely in the rate of organ donation. For example, 12% Germans donate, 100% Austrians, 98% Belgians and 28% Dutch. These results come as a surprise, because these countries have similar cultures and economies. It has been found that in the countries with the low-rate of organ donations the default option was not donating organs. When one wants to donate one’s organs one has to declare an explicit consent, e.g., while receiving a driving license. On the other hand, in the countries where almost everybody donates their organs, the default option was donating organs. The minor change in default option has an impact on

the big social issue. Moreover, most people declare that they want to donate organs, but when it comes to making an effort and declare consent not many do it. The small inconvenience (effort) connected with the registration was enough to deter potential donors⁴ (Thaler, & Sunstein, 2008, pp. 178, 179).

Thanks to default option the problems with *status quo* bias can be resolved. The example is the program “Save More Tomorrow” proposed by Benartzi and Thaler (2007). The goal was to increase the rate of savings for retirement. The question if people should save more is a normative question and this is not the reason for Thaler’s and Benartzi’s program. They point out that people on their own declare the need for more saving and they are dissatisfied with too low contributions they pay. Of course, deeds are more important than words. Finally, not many people who declare that they should save more change anything in their behaviour. The similar situation is with people who promise to quit smoking and exercise more. However, not many people promise to smoke more and exercise less. These kinds of statements (which can be perceived as precommitments/intentions) are treated by Thaler and Sunstein as the information that demonstrates that people are open for some strategies that can help them to achieve their goals (Thaler, & Sunstein, 2008, p. 107). This is a very controversial approach for many scholars. Later we will investigate the problems with these kinds of resolutions.

Generally, the automatic retirements plans assume a low rate of savings (2–3%) with the very conservative choice of investments. Even the saving rate of 6% is not going to change the situation dramatically. Because of the inertia effect, people very rarely change the rate of savings during the contract (Thaler & Sunstein, 2008, pp. 113, 114). Moreover, people spend little time to make a huge financial decision like retirement. It turns out that 58% of people spend less than an hour to fix the rate of contribution and make the investment decision (Benartzi, & Thaler, 2007).

For these reasons, the program “Save More Tomorrow” was introduced. It obligates the users to connect the rise of contribution for retirement with the rise in their salaries. Because of that people do not treat the rise of contribution as a loss because there is no decline in earnings. The program uses the inertia effect (automatic rise), but in this case inertia work in favour of more savings (Thaler, & Sunstein, 2008, p. 113). Most people who sign for the “Save More Tomorrow” stay there for four consecutive rises. After that, the growth was halted because

⁴ Thaler and Sunstein understand that the default option in the case of organ donations can be a sensitive issue. This is why they propose a compromise. When people extend their driver license, they will be obliged do declare their preferences for organ donation. An application for driver’s license extension will not be approved until one of the options is marked. Thanks to this, a lot of people can be saved and at the same time the freedom of choice is respected (Thaler, & Sunstein, 2008, p. 180).

the employees achieved the highest rate of saving in their retirement plans. After the success of the pilot program many firms started to use some planes based on the automatic indexation (Thaler, & Sunstein, 2008, p. 114).

Although Thaler and Sunstein support choice architecture, they understand that it is not a magic wand that resolves every problem. Choice architects are very useful, but only in some specific situations. They are needed when they can know what is the best for the agents. Choice architects can foresee what people want when they possess expert knowledge and the differences among particular tastes and preferences are not too large. (Almost everyone prefers, *ceteris paribus*, to be healthy than sick, live longer than shorter, etc. Sometimes we just know when we deal with Dr. Jekyll and when with Mr. Hyde.) Moreover, choice architects are useful when preferences can be easily discovered (e.g., when the government comes to the conclusion that people want a drug reimbursement program which offers low prices for drugs that are taken regularly). Choice architects will be helpful when one chooses a mortgage credit, but it is needless for a person who decides between Pepsi and Cola. To conclude, choice architects are useful when they have expert knowledge, choices are complicated, and differences in individual preferences are insignificant or they are easy to be assessed (Thaler, & Sunstein, 2008, pp. 247, 248).

7.3.3. Default option versus freedom

Both default option and choice architecture demonstrate how easy it is to influence people's choices. This is why some scholars object that it can result in manipulation and restriction of liberty (Gill, & Gill, 2012; Grüne-Yanoff, 2012; Mitchell, 2005; Qizilbash, 2012; Rebonato, 2014; Sugden, 2008b, 2018; White, 2013). Thaler and Sunstein respond to this criticism. They argue that in spite of default option the freedom of choice is preserved. How default option is used shows what libertarian paternalism is really about.⁵ The existence of default option as such is paternalistic. Finally, somebody decides for individuals what is appropriate. However, people can shift from default option to option that suits them the best. It provides a free choice and this should make libertarians satisfied in Thaler's and Sunstein's opinion (2008, p. 5). Some can argue that no default option should be provided and people must be left with an unconstrained free choice. However, as the example with the drug reimbursement system show us (six million people did not sign up)⁶ not everyone wants to have this choice.

⁵ It is important to underline the fact that Thaler and Sunstein are the supporters of free market and they do not want state to have more influence on people that it has right now. Their liberal agenda is visible when they gave support to a free market for an organ trade (Thaler, & Sunstein, 2008, pp. 174, 175).

⁶ The market oriented system that was introduced by George W. Bush (Thaler, & Sunstein, 2008, pp. 159-166).

As a result, these people are assigned to the random program. This is why the same question arises as in the case of the cafeteria. Would it not be better to give people who do not want to choose better conditions rather than random choice?

Thaler and Sunstein, even more importantly, argue that sometimes the state cannot remain neutral. It puts people in some context of choice which makes paternalism unavoidable. In the cases like organ donation the state creates the context of choice by establishing a default option as 'giving' or 'not giving organs'. Of course, the state can ask everyone individually and give driver licence after signing up for some option. However, it would generate enormous costs. From this perspective, default option is more effective. In general, Thaler's and Sunstein's argument about the ineditability of paternalism goes like this. If the state cannot remain neutral in some cases and it creates the context of choice for people, it is better for this context to be organized, transparent and beneficial for people rather than random. The problem with neutrality also occurs in the case of "Save More Tomorrow" in which the automatic signing to retirement system can be seen as a substantial interference in human choices. However, libertarian view that state's neutrality is achieved by not assigning people to any program is not true from the perspective of libertarian paternalism. It is a case, because not assigning people to any program is also a default option which influences their choices.

In general, libertarians argue that state should not exist or should be very limited (e.g., (Nozick, 1974; Rothbard, 1978)). In this regard, the criticism of libertarian paternalism concerns the justification of the state in general. Thaler and Sunstein do not respond to this criticism because they do not discuss whether the state should or not exist in theory. Libertarian paternalism is put into contemporary political and economic context in which the state exists and it influences people. Therefore, Thaler and Sunstein just want the state to work better (Thaler, & Sunstein, 2008, p. 237).

However, it is not hard to imagine the situations where default option will have too much influence on people and it will not leave enough freedom. The problem with default option is that its implementation can be very flexible. The change from the default option should be easy and it should not generate significant costs for people. However, we know that because of status quo bias, even small costs of changing the default option can be high enough to discourage people from choosing the option they really want. Thaler and Sunstein understand that there is no strict line between libertarian paternalism and direct intervention of the state. They argue that common sense is enough to decide in particular cases whether freedom of choice is preserved (Thaler, & Sunstein, 2008, p. 237). However, it is not always so clear. The example was the choice between OFE and ZUS, which Poles had to make. The default option was to be assigned to ZUS. If somebody wanted to put some of one's savings in OFE, one needed to send

the document (by e-mail or post). It was not very inconvenient, especially when we look at how important this decision was. However, one can imagine that more impediments connected with the red tape can result in a situation where not many people will change the default option. Therefore, the state can easily make the 'gate' in a system with default option narrower and it will restrict freedom of choice significantly (Ostapiuk, 2019c).

7.3.4. Some constraints on libertarian paternalism

In spite of the problems with the state and its potential influence on people, the state is the main choice architect in libertarian paternalism. As was indicated before, Thaler and Sunstein do not deal with the state in a theoretical context. They assume that the state exists and is unavoidable. Starting from this assumption they try to show how the state can work effectively. Nowadays, governments have a significant influence on people. Thaler and Sunstein want this influence to be more organized. For them, it should be based on soft paternalism.⁷

Thaler and Sunstein want the state to have as little influence on people's choices as it is possible. Although default option is an essential part of their approach, they think that it should be used only in specific cases. In general, the main task for the state is to make information more transparent for people. Thaler and Sunstein gave the example of credit services to indicate what should be the main role of the state in libertarian paternalism. The introduction of credit cards in the 1980s has changed the model of consumption in the USA. Banks abuse the cognitive biases of people which results in easiness and abstractness of spending money. The next big problem is the transparency of credits. Banks frequently make it harder for people to know the exact rates of interests. There are so many different kinds of interests that it is not comprehensible for many people. Thaler and Sunstein argue that banks should send the year reports in which all the rates are summarized. These reports have two goals. Firstly, users of credit cards could easily compare their charges with the charges in the competitive banks. Secondly, clients would have a better understanding of what they pay for (Thaler, & Sunstein, 2008, p. 143).

The next criticism of libertarian paternalism concerns the fact that cognitive biases do not apply to people only but they also concern the governments and choice architectures (e.g., (Glaeser, 2006; Qizilbash, 2012)). Thaler and Sunstein noticed these problems as well and they proposed the cost-benefit analysis as the resolution. Nowadays almost in every developed country, a public

⁷ With time their approach has gained more importance and now many countries try to implement it. In 2009 Sunstein took the position of the director of regulation in the USA. Thaler year later started to be the adviser on the special behavioural economics unit which was listened by the prime minister David Cameron.

institution needs to use the cost-benefit analysis. Thanks to it, choice architects can base their decisions on rationality and ‘system 2.’ They will be focused on facts and not on their own hunches (Sunstein, 2013a; Sunstein, & Thaler, 2003). Later we will see that not everybody agrees that cost-benefit analysis resolves the problems with cognitive biases concerning choice architects.

7.4. Libertarian paternalism in the light of capability approach

At the end of Chapter 6, it was mentioned that libertarian paternalism can be treated as the extension or some version of the capability approach. This understanding is possible for a few reasons. Firstly, both libertarian paternalism and capability approach have a negative attitude towards neoclassical economics. Both approaches especially criticize the revealed preference theory which assumes that people are rational and they choose what is the best for them. The supporters of libertarian paternalism and capability approach disagree with this assumption because preferences are malleable and people often do not choose what is the best for them. Secondly, in capability approach as well as in libertarian paternalism the focus is put not only on the achievement but also on the context of choice. It is especially important for the capability approach in which an opportunity criterion is the essential part. Also, in libertarian paternalism the opportunity to choose the other options is underlined as something fundamental. Even though it uses default option.

What makes possible to perceive libertarian paternalism as a framework similar to the capability approach is the Reason which is put on a pedestal by both approaches. It was spotted by the other scholars. Qizilbash (2009) argues that Sen’s formulation ‘reason to value’ has a similar role to the information and rationality criteria which are used in informed and rational desires accounts. Libertarian paternalism is also the approach that underlines the importance of rationality and information. Qizilbash argues that “on my reading, it is clear that – while they do not explicitly endorse this – they implicitly adopt a version of the informed or rational desire or preference view of welfare” (Qizilbash, 2009, p. 19). This understanding is possible, because Thaler and Sunstein frequently argue that they want people to make an ideally rational and informed choices which are not compromised by cognitive constraints which lead to lower well-being. They write that “in some cases individuals make inferior decisions in terms of their own welfare – decisions they would change if they had complete information, unlimited cognitive abilities, and no lack of self-control” (Sunstein, & Thaler, 2005, p. 177).

Despite the similarities, there are many points where these two approaches differ. The first point of discontent is the attitude towards paternalism. Thaler

and Sunstein embrace paternalism whereas the capability approach is against it. As was indicated in the previous chapter, Sen dismisses Sugden's argument that the capability approach can infringe freedom (Sen, 2006). Moreover, he responded to concerns made by Layard who cautioned against the danger of paternalism in the capability approach (Layard, 2005, p. 113; Sen, 2009, pp. 272–276). Also, Nussbaum (2000, pp. 51–59) argues against paternalism in her version of the capabilities approach. In general, Sen is against paternalism in his approach.

There are other differences. The capability approach is a much more pluralistic framework in which people can have different values. Libertarian paternalism rather tries to indicate the goals that ideally informed and rational agent would choose. It looks for the ultimate goals which most people want to pursue (e.g., longer, healthier life, more money *ceteris paribus*). Therefore, libertarian paternalism is more about searching for means by which people can achieve these uncontroversial goals. Whereas the capability approach is mostly about discussing different goals and this approach is reluctant to establish the goals which are good for everyone. Moreover, freedom has an intrinsic value for Sen and it is more important than well-being. Although Thaler and Sunstein underline the importance of freedom, it does not have an intrinsic value and it can be traded for more well-being. They also are satisfied with a less demanding version of freedom which is perceived from a negative perspective (negative freedom).

In general, the capability approach is a much more flexible approach than libertarian paternalism. Sen purposefully did not want to indicate the ultimate goals because it means being paternalistic. This is why he provides the concept of capabilities which is hard to pin down. However, his statement 'reason to value' opens the door to the approaches such as libertarian paternalism that want to indicate what values should be chosen by reason. Sen did not want to do it. This is why he probably would not like libertarian paternalism to be treated as the extension of his framework. However, it does not change the fact that libertarian paternalism can be perceived to some degree as the practical application of the capability approach.

7.5. The criticism of libertarian paternalism

7.5.1. Introduction

The most objected part of libertarian paternalism is the term 'libertarian paternalism' itself. Qizilbash (2009, p. 4) concisely summarizes what this term means:

This is, on their view, a form of 'paternalism' since it involves interfering with people's choices with a view to improving their welfare, while respecting freedom of choice. When human beings fall short in terms of either informed

or rational choice, the results of behavioural economics can help private institutions or the state improve their decision making and increase people's welfare.

Libertarian paternalism was proposed by Sunstein and Thaler (2003; Thaler, & Sunstein, 2008) but it is a part of a bigger approach supported by behavioural economics which is called 'soft paternalism.' Thaler's term gained the most recognition, because he is a pioneer in behavioural economics and he is one of its leading practitioners (Sugden, 2009a). At the same time as libertarian paternalism very similar idea emerged called "asymmetric paternalism" (Camerer, Issacharoff, Loewenstein, O'Donoghue, & Rabin, 2003). Loewenstein which was one of the people that introduced this approach worked with psychologists and they introduced a more sophisticated version of soft paternalism (Loewenstein, & Ubel, 2008). Even though there are many theoretical accounts of soft paternalism, in this chapter we are focused on the popular book *Nudge* written by Thaler and Sunstein (2008). Sugden (2009a, p. 366) underlines the importance of this book:

although *Nudge* is written in the breezy style of books targeted at business travellers, it can be read- and I shall read it- as a serious work of advocacy in support of a method of normative analysis that commands widespread support among behavioural economists.

We indicate three primary sources of criticisms of libertarian paternalism. Both criticism and answer to it are presented to understand whether libertarian paternalism can be considered as a significant contribution in terms of economics' explanatory and cognitive value as well as a practical application for people.

7.5.2. The concept of purified preferences. The problems with the inner agent

To begin with, it must be stated that Thaler and Sunstein think that the revealed preference theory does not indicate what is good for people. They argue for informed or rational approach in which it is possible to discover what is the best for people. For Sugden the problem with this view is the question who decides which desires are informed or rational and why they are good for people. This argument was used against the capability approach as well. Sugden and others are worried that some experts, philosophers or governments will decide instead of people what is good or bad. As a result, it will lead to the restriction of liberty. This argument goes like that:

If people knew the consequences of eating unhealthy food or smoking, they would realise that it is not good for them. Given that people are limited in rationality and imperfectly informed, society might then impose restrictions on smoking or on eating some sorts of food. That would be paternalistic. (Qizilbash, 2009, pp. 7, 8)

The problem for Sugden and others is that if we do not respect people's preferences, it will leave the room for abuse of freedom. Another problem for Sugden is that it is impossible to know these rational and well-informed preferences ("purified preferences"⁸). Even if scientists can define objectively what perfect rationality is, there might not be the way to determine what one would have chosen, if one had been fully rational (Sugden, 2009a). We might do not know that because people do not need to follow ideal rationality which is assumed by Thaler and Sunstein who believe that "inside every Human there is an Econ" (Sugden, 2009a, p. 370).⁹ Sugden argues against the concept of ideal econ. He and others write that "the inner agent is pictured as the locus of the identity of the human being and as the source of normative authority about its interests and goals" (Infante et al., 2016, p. 1). Sugden does not like the idea of inner agent and he criticizes this view explicitly elsewhere. He argues:

I believe, is that when economists (and indeed many philosophers, and perhaps even some psychologists) think about human agency, they find it hard to avoid using a mental model in which humans are ultimately rational beings. This model may recognize that humans can hold irrational beliefs and make irrational decisions, but at some deep level, irrationality is understood as the product of mistakes. These mistakes must be defined relative to some 'true' preferences – the preferences of the human individual's 'true self.' This is the model of the inner rational agent. We need to recognize that this model is pre-scientific. (Sugden, 2015, p. 598)

The proponents of libertarian paternalism believe in the inner agent which has purified preferences that should be followed, because they are the best for people. Thaler and Sunstein (2008, p. 5) argue for this position because "in many cases, individuals make pretty bad decisions – decisions that they would not have made if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control". Sugden indicates that the libertarian paternalism assumes that "what makes an individual better off 'as judged by himself' is defined by the preferences he would have revealed, had his decision-making not been affected by limitations of attention, information, cognitive ability or self-control" (Sugden, 2015, p. 583). This is why libertarian paternalism "treats context-dependant choices as the result of errors of reasoning" (Sugden, 2015, p. 583).

Sugden points out that to know what the ideal preferences are, we need to reconstruct "individuals' *latent preferences* by simulating what they would have chosen, had their reasoning not been subject to these errors" (Sugden, 2015, p. 583). This is what the process of preference purification means. Sugden argues

⁸ What a person would choose if it was not for cognitive biases (Hausman, 2012).

⁹ This objection can be somewhat controversial because Thaler is well aware of the fact that the ideal rationality is a normative, not a descriptive theory (Thaler, 1980).

that libertarian paternalism “can overcome the problem of context-dependence in actual choices only if, as Sunstein and Thaler implicitly assume is the case, the corresponding latent preferences are context-independent” (Sugden, 2015, p. 583).

The fact that preferences are context-dependent was demonstrated on the example of menu-dependence which was introduced by Sen (1997, p. 752, 2002). Sugden and others (Infante et al., 2016) refer to Hausman’s interpretation of the concept of preferences used in economics to points out this relativity. Hausman gives an example to show how the concept of preferences works. He indicates that “to say that Jill prefers x to y is to say that when Jill has thought about everything she takes to bear on how much she values x and y , Jill ranks x above y ” (Hausman, 2012, p. 34). Therefore, preference is comparative (x is ranked above y). This comparison with regard to value and the valuation is subjective (“how much she values” (Hausman, 2012, p. 34)). Moreover, it takes into consideration all aspects that the agents consider as significant to the comparison (“everything she takes to bear on”, p. 34). To sum up, a preference is a “total subjective comparative evaluation” (Hausman, 2012, p. 34). Hausman (2012, p. 35) argues that this definition “matches most of current practice” in economics.

Sugden (2015, p. 580) highlights the problems which are the result of the fact that preferences are context-dependent. He writes that:

Many economists and philosophers find the idea of respecting context dependent preferences problematic, either because there seems to be no good reason for thinking that such preferences are indicators of individual well-being, or more fundamentally, because the concept of respecting a person’s preferences is thought to be ill-defined unless those preferences satisfy minimal properties of internal consistency.

Sugden argues that the supporters of behavioural economics do not follow their premises. They do not want to conclude that preferences should not be respected at all and that experts should use their own judgments to decide on behalf of people about their well-being. Behavioural economists are reluctant to admit that because this position seems “unacceptably paternalistic” (Sugden, 2015, p. 580). Sugden perceives this as the inconsistency. If one assumes that context-dependent preferences should not be respected because of the problems, one should go all the way and call the whole concept of consumer’s sovereignty into question.

Sugden agrees with the findings of behavioural economics that people’s preferences are not coherent. However, he underlines that Thaler and Sunstein did not propose a strict method to find ‘purified preferences.’ They proposed only some vague hints on how to know what people’s real preferences are. The idea of purified preferences is essential in libertarian paternalism, because choice architects should nudge people towards choices that would be chosen if it was not for cognitive biases and lack of information (viz. purified preferences).

Thaler and Sunstein know that the state's nudges are paternalistic, but they are justified, because they are created to "make choosers better off, as judged by themselves" (Thaler, & Sunstein, 2008, p. 5) and this justification is repeated in the other places in the book (e.g., pp. 10, 12, 80). However, Thaler and Sunstein give very few insights on how we can discover those judgments (Sugden, 2009a). Because of the potential power that is given to the government in libertarian paternalism, it needs to give clear criteria for how to discover purified preferences.

For example, Thaler and Sunstein give a vague explanation for nudges against obesity, drinking, and smoking. To justify nudges, they provide statistics concerning health risks. They conclude that "with respect to diet, smoking, and drinking, people's current choices cannot reasonably be claimed to be the best means of promoting their well-being" (Thaler, & Sunstein, 2008, p. 7). The same method is used to justify nudges towards more saving. Thaler and Sunstein highlight that the rate of saving in the US is very low, and then they claim that "it seems clear that the costs of saving too little are greater than the costs of saving too much" (Thaler, & Sunstein, 2008, p. 106). Sugden argues that probably most people would agree with the statement that:

there are many people who would be well-advised to eat less saturated fat, drink less alcohol, stop smoking and save more. But that is not the point. What, according to their own account, Thaler and Sunstein have to show is that those people are making bad choices as judged by themselves. (Sugden, 2009a, p. 371)

To justify the argument that people really want to be nudged, Thaler and Sunstein occasionally refer to 'New Year's resolution test.' They ask "how many people vow to smoke more cigarettes, drink more martinis, or have more chocolate donuts in the morning next year?" (Thaler, & Sunstein, 2008, p. 73). In the case of saving they use the survey that shows that two-thirds of employees claim that their savings rate as 'too low' when only one percent characterize it as 'too high'. Thaler and Sunstein (2008, p. 107) argue that "such statements are not meaningless or random." Sugden agrees with that. However, he argues it is not enough, because we need a criterion that "could discriminate between the New Year's resolutions that many of us make without seriously expecting (or even trying) to keep and genuine personal commitments that fail only under intense psychological pressure" (Sugden, 2009a, p. 371). In Sugden's view, Thaler and Sunstein do not provide such a criterion.

An even greater problem than lack of precision is that Thaler and Sunstein do not seem to think that this precision is required (Sugden, 2009a). They agree that the New Year's resolution test has obvious limitations and that "people's actions may tell us more than their words" (Thaler, & Sunstein, 2008, p. 107). However, they argue that statements like "I should be saving (or dieting, or exercising) more" means that people "are open to a nudge" and "might even be grateful

for one” (Thaler, & Sunstein, 2008, p. 107). Therefore, Thaler and Sunstein do not claim that these statements give evidence that people want to be nudged but only that they might want to be nudged. This is the problem for Sugden who thinks that we need a precise criterion because nudges can have a major influence on individuals’ lives and they cannot be treated so lightly as it is done by Thaler and Sunstein (Sugden, 2009a).

The problem of the vagueness of the method also concerns the government and cost-benefit analysis. Thaler and Sunstein argue that planners should use where it is possible “a form of cost-benefit analysis” whose goal is “to measure the full ramifications of any design choice” (Sunstein, & Thaler, 2003, p. 1190). They understand that in many cases this method can be expensive and difficult. Therefore, it should be subjective and open-ended. Despite the relativity, the cost-benefit analysis can work. Thaler and Sunstein present the cost-benefit analysis on the example of automatic enrolment in savings plans. It should demonstrate that it will make some people better-off and some worse-off and that “a cost-benefit analysis would attempt to evaluate these gains and losses” (Sunstein, & Thaler, 2003, p. 1190). Thaler and Sunstein assume that this cost-benefit analysis will indicate that gains are larger than losses. They write that “in the abstract, it is reasonable to believe that the costs of having too little saved up for retirement are typically greater than the costs of having saved too much” (Sunstein, & Thaler, 2003, pp. 1190, 1191). However, they do not show how this calculation is done exactly (Sugden, 2008b).

Because of the problems with the cost-benefit analysis, Thaler and Sunstein argue that planners can just use “rules of thumb” to know what is in the best interest of people (Sunstein, & Thaler, 2003b, p. 1190). However, as Sugden (2008b, p. 233) points out:

it is hard to see how one can judge whether a rule of thumb provides a good approximation to the results of a full cost-benefit analysis unless one knows how, in principle, a cost-benefit analysis would work. The vagueness of Sunstein and Thaler’s descriptions of cost-benefit analysis strongly suggests that the difficulties they have perceived are conceptual rather than practical.

Sugden’s criticism does not mean that he is against the foundation on which libertarian paternalism is based upon. Sugden agrees that preferences are unstable and context-dependent (Sugden, 2008b, p. 227). In libertarian paternalism these findings are used to argue that the state can help people with achieving well-being. However, Sugden (2008b, p. 234) argues that “Sunstein and Thaler have not shown how the truth value of this belief can be assessed”. Sugden understands that Thaler and Sunstein can respond to this criticism by stating that paternalism is unavoidable and planners must act even if they do not have objective criteria for purified preferences (Sugden, 2008b). The truth of this claim will be assessed later.

The last but not least issue for Sugden is the fact of the normativity of informed and rational desire accounts. For Sugden it is impossible to treat this criterion as empirical. We cannot know without making normative claims what complete information or unlimited cognitive abilities are (Sugden, 2008b). Moreover, Sugden thinks that this normativity enables choice architects to justify any kind of intervention and it is a real danger of libertarian paternalism. Sugden makes these point explicit when he writes that:

The concepts of full attention, perfect information, unimpaired cognitive ability and complete self-control do not have objective definitions; they are inescapably normative. Just about any intervention that a paternalist sincerely judges to be in the individual's best interests can be justified in this way if the paternalist is allowed to define what counts as attention, information, cognitive ability and self-control. (Sugden, 2009a, p. 370)

Answer to the criticism

We can respond to Sugden's criticism by using the previous conclusions from this book. Hausman argumentation will be used as well (Hausman, 2016). To begin with, we need to recognize and agree with Sugden that rational desire account (purified preferences) is a normative conception. However, as was concluded in Chapter 6, normativity does not mean that something is unscientific or unworthy of study. Therefore, there is nothing wrong, in principle, with the fact that the 'rational desire account' is normative. Moreover, the revealed preference theory which is criticized by Thaler and Sunstein is a normative conception as well. (If one only describes what people do, it is not a normative claim. However, when one assumes that what people do is the best for them, then one makes a normative claim.)

Sugden argues that we do not know how and if we can ever know what purified preferences are. However, in Chapter 4, it was indicated that sometimes we know when we deal with Dr. Jekyll and when with Mr. Hyde. Of course, these kinds of statements are very vague, and Sugden is right when he criticizes Thaler and Sunstein for not providing the strict method to find what purified preferences are. However, one can use the findings from economics of happiness and behavioural economics to have some idea of what purified preferences are. Thanks to this knowledge, economists should be able to choose between the 'long-term human' (purified preferences) and 'short-term human' (revealed preferences). The enriched conception of multiple self can be used to justify Thaler's and Sugden's position. Also, the previous discussion about the weakness of will supports the idea of purified preferences. Of course, there is a problem between real commitments that people have and the others that are the result of social pressure. This is why the 'New Year's resolution test' cannot discover purified preferences in every case and we should take some of people's declaration with a pinch of salt (Bryan, Karlan, & Nelson, 2010; Carrera, Royer, Stehr, Sydnor, & Taubinsky, 2019;

Laibson, 2015; Sugden, 2018). However, it seems as a good starting point which can be improved by further statistical evidence. For example, we can discover if people's resolutions are real by looking at how many people actively try to quit smoking (precommitment strategies).

It is beyond doubt that the criticism expressed by Sugden is valuable. However, we need to look at the problems of purified preferences from a broader perspective. It is not enough to show the problems with purified preferences to dismiss this concept entirely. One needs to compare this approach with its competitor – revealed preference theory which is not without its flaws as was indicated through this book. It is true that the choice between these two competing approaches is normative. However, there is no other way and scientists sometimes must choose normatively between the competitive conceptions (Kuhn, 1962).

We need to write a few words in defence of the state and cost-benefit analysis. Sugden and many orthodox libertarians are afraid that libertarian paternalism gives too much power to the state, which will abuse people eventually. They use here 'slippery-slope argument.'¹⁰ To limit the influence of the state, Thaler and Sunstein endorse Rawls's 'publicity principle' which in its simplest form "bans government from selecting a policy that it would not be able to defend publicly to its own citizens" (Thaler, & Sunstein, 2008, p. 244). This rule of transparency of state's functioning is not enough for some. Qizilbash argues that "The 'publicity principle' only requires that the government be able to defend a proposed intervention, not that public reasoning would actually endorse the intervention through debate and discussion" (Qizilbash, 2009, p. 39). Qizilbash's unease with libertarian paternalism is understandable. However, this kind of criticism does not exclusively concern libertarian paternalism but, in general, it concerns the democratic system because it is hard to draw the line between the rules introduced thanks to public discussion and the rules imposed by politicians. In democracies there is a system of check and balances that should defend people from the force of the government. It works with various results but it seems the best that people can hope for.

When it comes to the cost-benefit analysis, Thaler and Sunstein acknowledge its problems and they propose the approach based on the 'rule of thumb' which is open-ended. It is a problem for Sugden who thinks that one needs to have precise criteria for how some method works. It does not necessarily need to be the case because sometimes we cannot have absolute criteria that should be obeyed in all cases. In social sciences, and especially in public policy, vagueness is sometimes unavoidable. It can even be an advantage, because we do not need

¹⁰ It is a type of logical fallacy. It suggests that taking a minor action will lead to extreme and sometimes ludicrous consequences.

to stick to one rule and we can be open-minded.¹¹ This is why an openness of cost-benefit analysis can be perceived as an advantage. Thaler and Sunstein argue that even if few people after the reflection and over time choose to opt out for savings plans, it indicates that people want to save more. This argument can be unconvincing for Sugden. However, the lack of precise criteria does not dismiss this approach, because sometimes we cannot provide strict criteria for some conceptions as was indicated in Chapter 5.

Now, we can examine the arguments made by Hausman who seems to support libertarian paternalism to some degree. He responded (2016) directly to the criticism from Sugden and rest (Infante et al., 2016). They perceived Hausman and the idea of purified preferences as the embodiment of the problems with the whole project of behavioural welfare economics in which the assumptions of revealed preferences are questioned. This criticism also concerns libertarian paternalism. Hausman (2016) indicates three kinds of criticisms presented by Sugden and rest (Infante et al., 2016).

1. The methodological criticism. It objects to the idea of context-independent preferences of the inner agent. Sugden and rest want to know how the inner agent forms its preferences. They also ask how it is possible that the inner agent has context-independent preferences. Moreover, they wonder if we can isolate a perfectly rational self. As was indicated before, Sugden and rest answer these questions negatively and they explain that the inner agent “is not a real human being whose thoughts and actions are governed by psychological mechanisms” (Infante et al., 2016, p. 13, 14). Hausman responds to these reservations by stating that “it seems to me that behavioral economists can deny that they are committed to any inner agent or to any process whereby they construct context-independent preferences” (Hausman, 2016, p. 27). Hausman expands this point by referring to the example of the cafeteria:

To suppose that there are truth conditions for the claim that an agent has a true or purified preference for fruit, when the agent’s choice between cake and fruit depends on their location in the cafeteria line, does not commit the behavioral economist to postulating the existence of an inner agent who is capable of weighing the various considerations bearing on the choice between fruit and cake and determining an overall preference. (Hausman, 2016, p. 27)

It means that we do not need a monumental ontological conception of the inner agent in which it is justified why people always prefer fruit over a cake. It could only imply that in this particular case an agent is concerned about her health and physical appearance and regrets sugar binges (Hausman, 2016, pp. 27, 28). We do not need to invoke any conception of purified preferences or inner agent to

¹¹ Over the last years situational ethics has gained a lot of importance precisely because of the lack of absolute rules (see (Doris, 1998)).

assume that if an agent were fully rational, she would prefer the fruit (Hausman, 2016, p. 28). Hausman indicates that “facts about the world provide evidence that the fruit promotes the agent’s health and appearance and features of the agent’s behavior and speech provide evidence that health and appearance matter more to the individual than the pleasures of sweets” (Hausman, 2016, p. 28).

The conclusion from Hausman’s argument is that we do not need the ontological conception of the inner agent. We can analyse case by case and decide what people want. Hausman is right, but it does not mean that Sugden is wrong. We need to have some criteria to know what the purified preferences are in a particular situation. These rules do not need to be absolute but we should have a reliable mechanism to know what people want. From the perspective of this book, observing people’s intentions and precommitment strategies is a good starting point to discover what people really want.

2. The epistemological criticism. It is similar to the methodological criticism. It also concerns the argument that it is impossible for behavioural economists to know what the purified preferences of inner agents are. In spite of being critical towards this idea, Sugden and others (Infante et al., 2016) admit that in some cases purified preferences can be known. They show the example of a consumers who fail to purchase electricity (an example of a homogeneous product) from the cheapest provider (Wilson, & Price, 2010). However, they treat this example as an exception. They want to know how it is possible for behavioural economists to find out whether the inner agent prefers fruit to cake when the choice depends on the placement of these goods. Sugden and rest understand that information about the past behaviours of people can indicate what they want. However, it seems that there is not much detailed evidence to be sure what purified preferences are. The more substantial problem for Sugden is that choice architects believe that they know what is the best for people. They assume that fruit is better for almost everyone and it satisfies purified preferences. The fact that some behavioural economists argue that their goal is to satisfy purified preferences of consumers and that they do not impose some arbitrary choices on people is a very dangerous practice for Sugden and others (Infante et al., 2016; Sugden, 2018). It is the case, because in the past when a government imposed some ‘best’ choices on people it was possible to discern that they are chosen arbitrary and individuals’ preferences did not matter. Libertarian paternalism is much trickier, because as a politician one can believe that one promotes somebody’s purified preferences and be happy about it. Hausman points out that it can be a problem, because in reality one does not know what purified preferences are, but one only assumes what they should be (Hausman, 2016, p. 28).

Although Hausman recognizes the problems with finding purified preferences, he argues that thanks to the experiments we sometimes can know what individuals’ purified preferences are. He argues that “for example, one can

study whether most people choose cake or fruit when the two desserts are placed in equally prominent positions and when their calorie and nutritional content is prominently displayed” (Hausman, 2016, p. 28). In the case of saving for retirement, choice architect can make an unbiased (taking account of people’s cognitive biases) presentation about retirement plans and see whether people participate in a plan or not. Hausman (2016, p. 28) concludes that “if almost all of them choose to enroll, then there is a reason to believe that automatically enrolling employees with an opt-out option will satisfy their purified preferences better than requiring employees to opt in.” The experiments which are designed to give people the opportunity to choose should satisfy Sugden to some degree, because he is a supporter of opportunity criterion (Sugden, 2004).

3. Implicit normative criticism. It concerns the possibility of invoking the purified preferences to limit or change the choices that agents make. As was indicated earlier, Sugden and others are worried that the state can use the concept of purified preferences to decide on behalf of people. Moreover, Sugden sees the inconsistency in libertarian paternalists’ reasoning which concerns the concept of purified preferences. He argues that when one dismisses the idea of revealed preferences which is based on preference satisfaction, one also needs to dismiss the idea of purified preferences. It is the case, because purified preferences are also based on preference satisfaction, but we search for the ideal preferences. Hausman also acknowledges the potential inconsistency, but he defends the idea of the limited use of purified preferences. He writes:

I argue that the satisfaction of preferences does not constitute well-being and that social policy should not aim to satisfy preferences, except in those special circumstances in which preference satisfaction is a reliable indicator of well-being. Preference satisfaction is a reliable indicator of well-being when three conditions are met: people are competent evaluators (and free of rational flaws), they are self-interested, and they have accurate relevant information. When these conditions are met, what people prefer will be a good (albeit imperfect) guide to what is good for them. (Hausman, 2016, p. 29)

Hausman points out that only when specific conditions are met, social policies can aim to satisfy people’s preferences. For Hausman the conception of purified preferences “is neutral among conceptions of wellbeing” (Hausman, 2016, p. 29), because it does not assume that what individuals do is always the best for them. His approach is a “pragmatic suggestion” (Hausman, 2016, p. 30), not the welfare theory that always gives the answers.

To conclude, the position of Sugden and Hausman seems to differ substantially. However, the differences are the result of the lens by which they look at libertarian paternalism and behavioural welfare economics. Hausman focuses on particular situations in which the state can help people. He argues that these situations must be analysed to know if nudges are justified. On the other hand, Sugden

is focused on the possible abuse and danger that come from the government that is using the conception of purified preferences. Both of them agree that the revealed preference theory is not ideal. However, they have very different attitudes towards libertarian paternalism. Hausman is more optimistic whereas Sugden is pessimistic when it comes to the concept of purified preference as a basis for welfare theory and policy.

7.5.3. Paternalism is not inevitable

One of the main reasons which is used to justify a paternalistic dimension of libertarian paternalism is that paternalism is inevitable. Sunstein and Thaler (2003, p. 1162) argue for libertarian paternalism in contrast to “the dogmatic anti-paternalism of numerous analysts of law, including many economists and economically oriented lawyers.” Sunstein and Thaler claim that the dogmatic anti-paternalism is based on one false assumption and two misconceptions. The false assumption asserts that “almost all people, almost all of the time, make choices that are in their best interest or at the very least are better, by their own lights, than the choices that would be made by third parties.” The first misconception is that “there are viable alternatives to paternalism.” The second misconception is that “paternalism always involves coercion” (Sunstein, & Thaler, 2003, pp. 1163–1165). We will deal with each assertion in order.

Sunstein and Thaler (2003, p. 1163) argue that the first assumption (people do what is the best for them) is “testable and false, indeed obviously false.” They support this claim by using the findings from psychology and behavioural economics (Thaler, & Sunstein, 2008). Sugden (2008b) and Qizilbash (2009) agree that people do not always make good choices for themselves. However, they think that Sunstein and Thaler fight with a straw man. Nowadays, no economist believes that people are ideally rational and make the best choices. It is true that in the 1980s neoclassical economists were very sceptical of behavioural economics, because many of them really believed in *homo economicus* (see (Thaler, 2015)). However, today no economist holds this view. This is why the description of mainstream economics presented by Thaler and Sunstein (2008) to make their point is not true anymore (Qizilbash, 2009; Sugden, 2008b).

Sugden and Qizilbash do not seem to be completely right. Although economists do not believe that people are ideally rational and make the best choice for themselves, these assumptions are still the basic conceptions that are taught in economic universities and they shape a way of thinking of young economists.¹² Moreover, many supporters of a free market willingly use the argument that people are rational and make the best deals. When it turns out that people do not

¹² The discussion about how economics should be taught see (Decker, Elsner, & Flechtner, 2018; Ratajczak, 2014).

make the best choices, it is argued that it is their own fault (a responsibility that comes with choice as a basis for capitalistic system). In the end, maybe nobody believes that people always make the best choices. However, the assertion that people make good deals most of the time and they should be held responsible for their actions is not controversial for most. This is why the claim of Thaler and Sunstein that the model of *homo economicus* is perceived as reality by economists is justified to some degree.

The first misconception is that ‘there are viable alternatives to paternalism.’ Thaler and Sunstein argue that the findings of behavioural economics prove that paternalism is inevitable and that any opposition to paternalism is a “literal nonstarter” (Thaler, & Sunstein, 2008, p. 11). Moreover, they argue that revealed preferences should not be respected, because people’s preferences are context-dependent and unstable. It implies that for any decision there is some choice architecture which nudges people in some direction. Therefore, Thaler and Sunstein ask if there is any reasonable alternative to choice architects who nudge people towards the choices that are the best for them (Thaler, & Sunstein, 2008, pp. 4–11). We can get back to the cafeteria example to understand this argument. There, Thaler and Sunstein argue that the only reasonable option is to arrange the food in a way that improves a consumer’s well-being (from fruits to cakes). Sunstein and Thaler admit that this is a paternalistic behaviour. However, it is not a problem, because there are no alternatives to paternalism (consumer’s choice is always influenced by something because of context-dependent preferences). This is why no anti-paternalist position makes sense. If that is the case, then the only task for libertarians is to provide freedom of choice. Sunstein and Thaler (2003, p. 1102) make this point explicit:

The central point is that effects on individual choices are often unavoidable. Of course it is usually good not to block choices, and we do not mean to defend non-libertarian paternalism here. But in an important respect the antipaternalist position is incoherent, simply because there is no way to avoid effects on behaviour and choices. The task for the committed libertarian is, in the midst of such effects, to preserve freedom of choice.

Thaler and Sunstein recognize that both firms and governments can be choice architects and they can nudge people (e.g., the influence of employers in the case of retirement savings). However, it does not mean that both of them are paternalistic in the same sense. Qizilbash (2009, p. 24) points this out:

But even if people or organizations make (or have to make) those choices, they do not always do so with a view to improving our lives or welfare. They may do so with a view to pursuing their own interests or profits.

The difference between governments and firms is in the motivation why they nudge people. Firms want to earn money and not help people. Therefore, a market system is presented by Sugden as the opposition towards paternalism (see (Sugden, 2008b)).

In general, Sugden (2009a, 2018) thinks that libertarian paternalism is so popular because it appeals to anti-market sensibility. He is a supporter of market and he refers to Hayek's idea that "market is a spontaneous order structured by the seeking of mutual advantage" (Sugden, 2008b, p. 239). He argues (2008b) that mutual advantage exists on the market even if people's preferences are incoherent. Moreover, Sugden argues that market gives people what they really want. This is why he prefers a free market over libertarian paternalism:

Speaking for myself, and without wanting to claim that all forms of market regulation are undesirable, I welcome the fact that cafeterias are normally operated according to the principle of mutual advantage. I want it to be the case that they try to offer me products that I want to buy. I want their cakes to look attractive, and to be presented in ways that stimulate my appetite. It is not that I am a paragon of informed desire, acting on complete information with unlimited cognitive abilities and no lack of willpower. It is just that I would rather have my willpower challenged by tempting cakes than license cafeteria managers to compromise on the attractiveness of their products so as to steer me towards the ones that they think best for me. (Sugden, 2008b, p. 249)

Sugden understands that it is his personal opinion and that there are the others who may want to be steered towards the right choices. He writes:

No doubt there are other people who recognise that they have real problems of dietary self-control, and want to be steered away from unhealthy food. I am not claiming that it is wrong for public policy to respond to this want, if enough people feel it. (Sugden, 2008b, p. 247)

He is also not an unconditional supporter of a free market and he is in favour of some regulations that soft paternalism advocate. For example, he advocates the "standardised labelling of food products, standardised tariff structures to facilitate price comparisons, or mandatory cooling-off periods before sales of financial products are confirmed" (Sugden, 2008b, p. 247).

The main problem for Sugden is not libertarian paternalism itself, but the scale of the claim that its supporters have made. According to him, even if the revealed preferences theory does not work, it does not mean that a policy advocated in libertarian paternalism does not restrict freedom. These interferences limit a market mechanism that could work for people's benefits in normal conditions. Sugden ended his paper by dismissing Thaler's and Sugden's claim that paternalism is inevitable. He writes:

It is not enough to show that consumers' preferences are incoherent and to assert that arguments against paternalism have thereby become non-starters. The question that this paper has addressed is not whether paternalism is good or bad. It is whether, given the findings of behavioural economics, there are viable alternatives to paternalism. I have argued that the answer to this question is 'Yes'. There is a viable alternative to paternalism. It is what it has always been – the market. (Sugden, 2008b, p. 247)

Answer to the criticism

Sugden seems right in his suspicion that many behavioural economists are not especially fond of free market. In Chapter 3 in the context of subjective well-being it was presented that it is possible for scientists to abuse their power and to decide on behalf of people what is the best for them. A free market in many ways secures that no third wise party will decide what is good for individuals. This is why the unease of Sugden is justified. The state can use the findings of behavioural economics (e.g., context-dependent preferences) to argue that people need to eat healthily or drink less alcohol. It can lead to the abuse of freedom. However, it does not necessarily imply that we need to dismiss libertarian paternalism. We can limit and constrain the potential scope and influence of this approach.

In some way, it has been already done. Although Sugden is right to worry, it seems that he goes too far with his concerns. For Thaler and Sunstein a free market is still the most important mechanism that copes with people's preferences. The role of the state is limited. It should only influence people's choices in specific situations when the state can discover what most people want (not in the case of buying soda, but when people choose saving plans for retirement). Thaler and Sunstein do not force one to choose either free market or libertarian paternalism (identified with the state). It is crucial to remember that there are different levels of market intervention.

Sugden highlights the positive influence of a free market on people (mutual advantage),¹³ which is a true and important insight. However, one cannot forget that a free market does not work for the benefit of people in every situation and it does not always improve their well-being (Ostapiuk, 2018, 2019c). This point will not be elaborated further because the discussion on the advantages and disadvantages of market from the above-mentioned perspective goes far beyond the subject matter of this chapter (more about this discussion see in: (Frey, 2010, 2018; Friedman, 1962; Friedman M., & Friedman R., 1980; Galbraith, 1998; Hayek, 1978; Nozick, 1974; Offer, 2006; Scitovsky, 1976; Sugden, 2010, 2018)).

7.5.4. Libertarian paternalism does not provide freedom

The second misconception indicated by Sunstein and Thaler is that "paternalism always involves coercion" (Sunstein, & Thaler, 2003, p. 1165). As was presented on the example of the cafeteria, an interference in people's choices is inevitable. Therefore, paternalism is inevitable as well. However, Thaler and Sunstein

¹³ Bruni and Sugden (2013) argue that mutual benefit counts as a virtue for the market participants.

added the libertarian part to their conception, because interference in people's preferences is only justified when it does not limit the freedom to choose. Qizilbash (2009, p. 19) points out that "this makes the paternalism they are discussing 'libertarian' in the sense that it is a form of paternalism which should not be objectionable from the point of view of libertarians".

Thaler and Sunstein argue that their conception remains paternalistic even though it does not limit choice. The idea that paternalism needs coercion is, as was mentioned earlier, a 'misconception' (Sunstein, & Thaler, 2005, p. 180). For Thaler and Sunstein, paternalism implies that we want to increase people's well-being by helping them to choose better. In the cafeteria example, the arrangement of products in the way to make people better off is what makes it paternalistic. However, in their opinion libertarians should not object to that because there is no restriction of freedom. This definition presented by Thaler and Sunstein is criticized by Qizilbash who writes:

But the thought that 'paternalism does not require coercion' does not seem to be a 'misconception' about paternalism, but merely a decision about how to define it. Sunstein and Thaler make that decision, in part, to be able to advance the idea of 'libertarian paternalism' as a coherent concept. (Qizilbash, 2009, p. 24)

This semantics struggle is important, because by creating a specific definition of libertarianism, Thaler and Sunstein can argue that their conception should satisfy libertarians. This is not the case, because libertarians have a very specific understanding of what freedom means. For them freedom has an intrinsic value, which is not how freedom is perceived in libertarian paternalism. For libertarians Thaler and Sunstein add freedom dimension to their conception only to curb the potential criticism. This is why Thaler and Sunstein introduced the opt-out clauses in their approach. It is, however, not enough for libertarians who assign an intrinsic value to freedom. This is why Qizilbash criticizes libertarian paternalism. He argues that "the chief concern is not to advance autonomy or opportunity, but to improve decision making with a view to improving people's lives and increasing welfare while protecting free choice" (Qizilbash, 2011, p. 38).

Thaler and Sunstein used Mill's idea in which autonomy means that people should be able to make their own mistakes to argue that libertarian paternalism preserves freedom (Sunstein, & Thaler, 2005, p. 181). In the end, libertarian paternalism lets people make mistakes by the opt-out clauses. People can decide to eat only cakes and not save for retirement. Because of that, Thaler and Sunstein think that their approach should not be dismissed by libertarians. Qizilbash agrees that Sunstein and Thaler do not block choice and to that degree their approach does not undermine autonomy (Qizilbash, 2009, p. 23). However, libertarians cannot agree to the fact that the state actively tries to interfere with human preferences even if it wants to improve their well-being.

To understand that improving well-being cannot justify the interference with people's autonomy, we can refer to the 'prudential value theory' presented by Griffin (1986, p. 67):

one element of agency is deciding for oneself. Even if I constantly made a mess of my life, even if you could do better if you took charge, I would not let you do it. Autonomy has a value of its own.

This is the core of libertarians' arguments. Liberty has an intrinsic value and even if one suffers or is miserable, nobody can take charge of one's decisions. Libertarians argue that by nudging individuals the state takes charge of people's lives even if they have opt-out clauses (Gill, & Gill, 2012; Grüne-Yanoff, 2012; Mitchell, 2005; Rebonato, 2014; Sugden, 2008b; White, 2013).

It seems that there is no common ground between libertarians and nudges. Libertarian paternalism perceives people to be weak and in need of help. Even if this image is real (and findings of behavioural economics support this view), many libertarians will not accept it, because they cannot do it. Their whole worldview is based on the assumption that, in the end, people can and should decide on their own. Sometimes they do it better, sometimes worse, but it is demanded from people to take responsibility for their actions, because it is a part of human dignity. Therefore, for Sugden and Qizilbash the term 'libertarian paternalism' is an oxymoron even if Thaler and Sunstein try to present it as a coherent conception.

Answer to the criticism

The discussion on freedom is one of the greatest and most debated issues in science and the author will not even try to resolve this extremely intricate issue. The arguments presented in this section are directly connected with libertarian paternalism and should be perceived from this perspective only.

Let us start with the argument that concerns the infantilization of people. In libertarian paternalism, it is assumed that people are weak because of cognitive biases. This perception of people attracts much criticism. The critics pointed out that this approach infantilizes people because the state needs to decide on behalf of them; and freedom of choice is precisely what gives people dignity (Grüne-Yanoff, 2012; Mitchell, 2005; Pettit, 1996; Rebonato, 2014). It is a misunderstanding of Thaler's and Sunstein's position for whom the default option is not a basic tool that the state uses. The state can also increase people's freedom. For example, sometimes people do not want to make their own decisions and they rely on other people, institutions, or rules. In libertarian paternalism the state can force people to make an independent decision. Although it is a paternalistic action, it can increase people's freedom. (e.g., Bush's drug reimbursement program (Thaler, & Sunstein, 2008, pp. 159–166)).

The argument that the government in libertarian paternalism infantilizes people seems to be an exaggeration. This argument is not raised by the opponents of libertarian paternalism when in the face of some troubles people are looking for help from specialists such as doctor, mechanic, or lawyer. In libertarian paternalism, the state does not help people in every decision but only in specific ones. We do not see infantilization of people when the state helps them to choose among retirement portfolios if people do not have any idea about finance. From this perspective, the state does not prohibit people who have the knowledge and want to invest their all retirement money in the Romanian firm from the high tech industry. The main goal of libertarian paternalism is to provide warning signals for untutored customers. As it is done in the case of ski slopes where the inexperienced skiers are warned when the route is difficult (Thaler, & Sunstein, 2008, p. 241). The fear of ‘slippery slope argument’ in which it is argued that the state by helping people will make them unable to make any decisions independently in the future seems to be an exaggeration. We can compare this fear to one of the characters in *Phaedo* (Plato, 2013) who argues that the discovery of writing will result in the extinction of memory because people will not need it anymore.

Of course, it does not mean that we should not scrutinize nudges carefully. Sugden is right to warn about default option and the fact that it can have too much influence on people and it can prohibit people from following their preferences. Even if people have the opt-out option, they can be too lazy to use it. However, it can indicate that they did want to have freedom primarily. From this perspective, it does not matter whether people make decisions in libertarian paternalism or in a free market because in both systems people do not really exercise their freedom. However, at least thanks to the default option in libertarian paternalism people have higher well-being.

The next argument against the libertarian perception of freedom is the fact that freedom is not always good for people and it does not make their lives better. Schwartz (2004) argues that having too many choices makes people unhappy, because after a decision people are not sure if they have chosen the best option and it can lead to regret. Schwartz’s general conclusion is that people want freedom, but not too much freedom. Schwartz also argues that freedom is precisely connected with self-determination. People are taught that they are responsible for their actions. Both the successes and failures assess people as capable or incapable human beings. Schwartz indicates the problem with that view:

The more we are allowed to be the masters of our fates in one domain of life after another, the more we expect to be. Education is expected to be stimulating and useful. Work is supposed to be exciting, socially valuable, and remunerative. Spouses are supposed to be sexually, emotionally, and intellectually stimulating and also loyal and comforting. Friends are supposed to be fun to be with and devoted. Children are supposed to be beautiful, smart,

affectionate, obedient, and independent. Everything we buy is supposed to be the best of its kind. With all the choice available, people should never have to settle for things that are just good enough. In short, life is supposed to be perfect. Excessive emphasis on self-determination has, I believe, contributed to these unrealistic expectations. (Schwartz, 2000, p. 85)

The problem with that view is that people are not always in control. For example, some research demonstrates that people can do little about their weight (Seligman, 1994). Despite this fact, the contemporary western culture preaches that obesity is a matter of choice and personal responsibility. Nowadays, people in western countries are taught and believe that everything is in their control and power. People forget about fate and that there exist things that are not in their control which was indicated, for example, by ancient Greeks. In some way, people in western world started to believe in Sartre's idea that they are responsible for all of their actions (Sartre, 2007). In general, taking responsibility for actions is desirable and beneficial because people can improve and change their unwise behaviours. However, when one is not in control and one cannot change something, the idea of self-determination can make one only feels bad about oneself and it does not make any good for anyone.

The idea of learning on mistakes and taking responsibility for actions is essential for liberals. Hayek (1978, p. 18) argues that "liberty does not mean all good things". It means that freedom and making mistakes are intertwined. Moreover, he states that "liberty and responsibility are inseparable" (Hayek, 1978, p. 71). The supporters of libertarian paternalism do not disagree with the importance of responsibility. However, they point out the differences between choices. If choices are frequent and prosaic, people will learn on their own how to choose (e.g., buying bread that suits one's taste). The completely different situation is when the choice is rare and it has a substantial influence on life. There is no time to learn when one chooses a house or a mortgage credit once for several dozen years. In these kinds of decisions, choice architects can be helpful because the costs of a bad decision can be severe. As Thaler and Sunstein (2008, p. 241) point out:

We do not believe that children should learn the dangers of swimming pools by falling in and hoping for the best. Should pedestrians in London get hit by a double-decker bus to teach them to 'look right'? Isn't a reminder on the sidewalk better?

In the case of a danger even libertarians allow for the restriction of people's freedom because it is assumed that people in general do not want to die or to get harmed. More importantly in the case of danger people's freedom to live is endangered. The most well-known example that justifies the restriction of liberty was presented by Mill who wrote:

Again, it is a proper office of public authority to guard against accidents. If either a public officer or any one else saw a person attempting to cross

a bridge which had been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back, without any real infringement of his liberty; for liberty consists in doing what one desires, and he does not desire to fall into the river. (Mill, 2003, p. 165)

Because Mill was a supporter of freedom he immediately underlined the limited extent of freedom's restriction:

Nevertheless, when there is not a certainty, but only a danger of mischief, no one but the person himself can judge of the sufficiency of the motive which may prompt him to incur the risk: in this case, therefore, (unless he is a child, or delirious, or in some state of excitement or absorption incompatible with the full use of the reflecting faculty) he ought, I conceive, to be only warned of the danger; not forcibly prevented from exposing himself to it. (Mill, 2003, p. 165)¹⁴

In this context it seems justified to perceive libertarian paternalism as a way of warning. Thanks to the opt-out option it does not force one to something if one really does not want it. Of course, the term warning is vague. It is an open question to what extent and how we can warn people. The limits depend on the values with which one comes to the table (being a libertarian or supporter of state intervention). For libertarians default option is not acceptable but, for example, better information put on products which improves the transparency of information can be justified.¹⁵ Again, we can use Mill to justify these kinds of interventions:

Such a precaution, for example, as that of labelling the drug with some word expressive of its dangerous character, may be enforced without violation of liberty: the buyer cannot wish not to know that the thing he possesses has poisonous qualities. (Mill, 2003, p. 165)

As we can see, moderate nudges can be accepted by libertarians. However, most of such nudges are only interim measures which do not change people. They will not be able to learn on their own mistakes, and they will need nudges for the rest of their lives. Parents advise children because they believe that in the future children will be able to make independent choices. It seems that libertarian paternalism does not give this hope.

Let us investigate further the connection between individuals' moral responsibility for their actions and freedom. Fromm was one of the most influential critics of the assumption that this connection is always beneficial to people. He argued that capitalism does not offer any personal identity for people as it was in Feudalism when people knew exactly what their purpose and position in

¹⁴ Further discussion of Mill's position see (Dworkin, 1972).

¹⁵ Therefore, in the case of cafeteria, we could add the placards with information about the nutritional value of each food instead of manipulating the arrangement of food (Hausman, & Welch, 2010, p. 129).

life were. Capitalism resulted in people's lack of security and loneliness. Fromm argues that these are the strongest fears and people try to escape from them. Frequently, it is an escape from self by becoming a part of some larger, more powerful entity (e.g., God, nation or consciousness). Being a part of some entity, people participate in its power and glory. However, people need to surrender their freedom in exchange for security and proud that are guaranteed by a powerful entity like nation or God. In his book *Escape from freedom* (1994) Fromm points out how weak people are and how often they do not want freedom. This is the result of the fact that freedom of choice means responsibility and negative consequences. People escape from freedom, because they do not want to be responsible and they do not want to deal with the negative consequences of their actions.¹⁶

People frequently escape from freedom. Earlier the example of precommitment strategies was presented to prove that point. Some libertarians can argue that in the case of precommitment strategies people choose on their own to limit freedom. However, other people or firms may help people with sticking to precommitment strategies. For example, people can use firms as tools to help them quit smoking. People can make an agreement with some company that if the result of the test shows that they smoke, then they will need to pay the company money, and it should motivate them to quit smoking. Moreover, many people do not want to have choices and this preference can be treated as the default option. This is why many smokers were happy after the state introduced the bill that prohibited smoking in pubs, restaurants and the other public places (*Sondaż: Polacy zadowoleni...*, 2010). Generally one may say that when there is no choice, the likelihood of falling into temptation becomes smaller for most people and it is easier for them to stick to the commitment.

The last argument against treating freedom as intrinsic value refers to our discussion of 'basic' and 'nonbasic' values (Sen, 1970). When we treat freedom as a 'nonbasic' value, no discussion is possible. Moreover, it leads to dogmatism, which is often the case with the hard-core libertarians who do not want to compare freedom with the other values like well-being. Freedom should be fundamental, but one cannot forget that people frequently decide to have more well-being instead of more freedom. One needs to acknowledge and respect this choice. Intrinsic value does not need to mean unrespectable of other things. One can assign an intrinsic value to something, but at the same time one can compare this value with the other values and the comparison does not need to be all or nothing.

We can find this pragmatic understanding of freedom in libertarian paternalism. Thaler and Sunstein think that freedom has an intrinsic value which should be

¹⁶ Milgram's experiment is the best example of how people pass freedom to the authorities (Milgram, 1963).

respected. However, it can be compared and exchanged with the other intrinsic values like well-being. Sugden sees this relativity as the flaw because it can lead to abuse of freedom. In this book, we have been trying to justify a view that there are no absolute values and we need to learn how to live in the world without objective Truth. Therefore, relativity and pragmatism of libertarian paternalism do not necessarily need to be perceived as a flaw but as an advantage.

It can be seen as an advantage because there is no absolute freedom which is cherished by libertarians. The idea of freedom is not as uncontroversial as many of them think. We even do not know what freedom really means. Is it negative freedom or a positive one? Moreover, when one looks back at the discussion on self and personal identity, one can see that the conception of self and consequently freedom is not as straightforward as it seems at first glance.

Thaler and Sunstein propose, although not explicitly, their own version of freedom. The freedom which is connected with reason. They look for purified preferences which can be discovered thanks to deliberation and rationality. This argument is very close to the idea of 'long-term human.' Moreover, it can be supported by the arguments about the links between reason and freedom, which were presented in Chapter 6. One of its most important conclusions is that, in the end, people should be able to live the life that they have "reason to value" (Sen, 1999, p. 18).

7.6. Conclusions

The discussion about libertarian paternalism is necessary. Over the last years, many countries have been trying to implement some proposals presented by Thaler and Sunstein (2008). Therefore, libertarian paternalism started to have a significant impact on people's lives. The idea of libertarian paternalism is very controversial and it triggers a lot of objections and reservations in academia and outside. It comes as a no surprise in the light of the fact that such fundamental issues like freedom and the legitimacy of state interference are discussed in libertarian paternalism.

Much criticism of libertarian paternalism is valuable. We should be careful and scrutinize how much power is given to the state. Nudges cannot be taken lightly. We need to know the limits of their influence on people. Moreover, it is important to define the limits of libertarian paternalism better. Some flexibility is justified, but we need to have the criteria for purified preferences to know in which situations the state can interfere with revealed preferences.

From a philosophical perspective, it is crucial to perceive the claims made by libertarian paternalism moderately. The rational approach (purified preferences) is not an all-encompassing ontological theory of the inner agent. It is a proposition

which does not make such high claims. We should be conscious that libertarian paternalism is mostly of a pragmatic character. It is not a magic wand that will help solve or mitigate every problem. It can be used only in very specific cases which are indicated by Thaler and Sunstein (2008). Moreover, the main focus of this approach should be put on the transparency of information which helps people make a choice. Default option is very useful but it should only be used when the other forms of nudges such as transparency of information do not work properly.

Thaler and Sunstein do not provide sufficient philosophical background for their conception and because of that they were criticized. Even if pragmatic, their conception should have better defined rules and scope because otherwise it is vague and it attracts criticism. The previous chapters of this book can be used to provide some methodological grounding to libertarian paternalism. It was presented how we can find purified preferences and why they must be differentiated from revealed preferences (multiple selves and precommitment strategies). Although in some situations they will indicate what people really want, these conceptions will not provide us with an ironclad theory of purified preferences. However, when we look at the discussion on relativity in science it seems justified to argue that the libertarian paternalism' lack of formalistic structure does not legitimize its dismissal. In the end, everything is in the eye of the beholder. Some people can be in favour of libertarian paternalism and some are not. However, it seems that this conception can be defended against a big chunk of criticisms.

It is noticeable that the very term 'libertarian paternalism' attracts a lot of criticism. It is the result of the fact that paternalism and libertarianism can be defined differently and used for different purposes. The change of the term 'libertarian paternalism' is problematic, because it started to be popular and many papers were written using this term. Therefore, it is only suggested that the better (less controversial) term for Thaler's and Sunstein's idea would be 'soft paternalism' (popularized by Feinberg (1986, p. 12)) which was the term used before Thaler and Sunstein's book has gained popularity.

Moreover, it is important to underline that libertarian paternalism is in line with the argument presented in Chapter 6, where reason is identified with freedom. Libertarian paternalism supports and provides the practical implication to the main idea of this book that we should support 'long-term human,' because it identifies well-being with reason and it treats 'short-term human' (human with cognitive biases) as something that should be overcome. However, it must be added that the capability approach with the wider perception of agency and reason is much more in line with 'long-term human.' This is the case, because the other values rather than only ideal individual rationality are taken into consideration as it takes place in libertarian paternalism.

Conclusions

Chapter 1 starts with an analysis of the origins of value-free economics. It is demonstrated that logical positivism had a substantial influence on value-free economics. Because of logical positivism Robbins dismissed ethics from economics and he popularized a distinction between positive and normative economics. Moreover, the crave for objectivity caused that economics dismissed psychology and it ceased to discuss human goals and motivations. Later on, Samuelson argued that mathematics was a common language of economics and he tried to mathematically formalize economics which resulted, among other things, in the revealed preference theory. Moreover, it is indicated why strict normative/positive dichotomy is not justified from the epistemological perspective of the contemporary philosophy of science. The conclusion is that this distinction should be 'disinflated' (Putnam, 2002, p. 19). Moreover, positive economics contains normative elements.

The main goal of **Chapter 2** was the analysis of Becker's economic approach from the perspective of 'scientific research programme' (Lakatos, 1980). The 'hard core' of economic approach was indicated (maximization of utility and instrumental rationality). It was presented how the 'protective belt' was created around these assumptions which, consequently, enabled economists to absorb every kind of criticism from the other social sciences. The author made the analogy between the economic approach and value-free economics to demonstrate how axiomatic assumptions of value-free economics have led economists to overconfidence in explanatory strength of purely positive economics. Consequently, it has led to the dismissal of normative approaches in economics. Moreover, it is presented why the economic approach and the axiomatic assumptions of value-free economics have led to the tautology. It is argued that the economic approach is a degenerative programme, because new auxiliary hypotheses are not proposed to deal with new facts better but only to defend the 'hard core.' The main conclusion from **Chapter 2** is that value-free economics should not be built on the axiomatic assumptions. The paradigm must change because of the problems that are the result of the economists' reliance on axiomatic assumptions.

In **Chapter 3** the author investigates whether economics of happiness with the conception of subjective well-being can lead to the 'paradigm shift.' Although it is better at depicting and predicting people's behaviour than revealed preference theory, it creates problems on its own. The conclusion that we cannot assess

utility to find out what is best for people is well-known in the literature about happiness. The main task of this chapter was to compare revealed preference theory with different happiness approaches. The focus was put on the criticisms of revealed preference theory. Thanks to economics of happiness and positive psychology we know that revealed preferences are not always a good indicator of well-being. Individuals frequently make mistakes and they do not choose things which are the best for them. Despite the problems with revealed preferences theory, it was presented that economics of happiness has problems with subjective well-being. Firstly, malleability of preferences. As the result, establishing subjective happiness as the basis for well-being is problematic because it can change relatively easily. What is the most problematic issue here is that people can adjust to even dire conditions and feel happy. Secondly, there are other values than happiness which are not taken into consideration in subjective well-being measure (SWB). It can be very problematic, especially when people choose between freedom and happiness. Not distinguishing between values like happiness and freedom seems to overlook the motivations which strongly influence individuals' choices. Therefore, we need to introduce the idea of values and goals which are independent of revealed preferences where utility is maximized by definition. The possibility of that is provided by objective approaches towards well-being (e.g., the capability approach).

If we know that revealed preference theory's assumption (people do what is the best for them) is not a positive conception and it leads to many problems, then it seems justified for economists to choose normatively what is the best for people and to indicate goals which are independent of individuals' revealed choices. In **Chapter 4** the author used the conception of multiple-self to address this issue. It is especially important in the case of choices over time (different preferences). I tried to provide a better framework for the original conception of multiple self. I used for this goal the conclusions and insights from the philosophical discussion on personal identity.

One of the main goals of this chapter was to connect philosophical conception of intention with precommitment conception used in economics. I expanded Schelling's idea and the framework was proposed in which we can explore what individuals' real preferences are. The analysis was done in the context of weakness of will. It was presented why value-free economics cannot admit its existence. I showed a similar approach taken in philosophy (Hare, Plato) and it was demonstrated that the conception of weakness of will explains many phenomena that cannot be explained by value-free economics.

I used the ideas of such philosophers as Dworkin and Frankfurt to support the claim that reason ('second-order desires' (Frankfurt, 1971), 'meta-rankings' (Sen, 1977)) should indicate individuals' real preferences. Kantian approach was used to justify the primacy of reason. Moreover, this approach was used to show the

limitations of ‘belief-desire account’ introduced by Hume and which is used in value-free economics. The conclusion is that we cannot embrace reason based on categorical imperative within the revealed preferences’ framework.

Chapter 5 concerns philosophy of science. The purpose of that philosophical discussion was to highlight that absolute Truth cannot be found in science. This fact is well known in philosophy but not in economics. This is why it is interesting to analyse value-free economics from the perspective of the contemporary philosophy of science. It was demonstrated why value-free economics is perceived as objective by economists and why it is not true. Next, Weber’s insights were used to argue that we cannot escape from values in science. This analysis can be valuable, because Weber is frequently misinterpreted in economics as a person who preached value-free science. It was demonstrated why this misunderstanding happened. Moreover, it was indicated that economists cannot stick only to axioms and for the existence of any social science it is unavoidable to have some ‘methodological value judgments’ (Weber, 1949).

Chapter 6 was an analysis of the capability approach. I tried to explain how the capability approach works from an economic perspective and whether (or to what extent) it is better in explaining the complexity of human behaviour (e.g., values, motivations) than the revealed preference theory and the other approaches (economics of happiness). This analysis was conducted to find out whether Sen’s flexible framework can explain what people should do in order to increase their freedom and well-being. The analysis of the capability approach was also used to support and strengthen my conception of ‘long-term human’ (Ostapiuk, 2019b).

Chapter 7 provides some philosophical background to libertarian paternalism, which has not been done sufficiently by the founders of this approach. The previous insights from this book were used to justify Thaler’s and Sunstein’s claims. In particular, the discussion on weakens of will may help us to explore what individuals’ real preferences are. Thaler and Sunstein have not presented a convincing method to do that. The conceptions of intentions and precommitment strategy were proposed as the methods by which one can answer that question.

Thanks to a philosophical background, it was possible to respond to the criticisms of libertarian paternalism and even to the broader criticisms of welfare behavioural economics. However, providing the philosophical justification for libertarian paternalism does not mean that I am uncritical of this idea. Libertarian paternalism is not the best term because of its controversial nature. Instead, the term ‘soft paternalism’ (Feinberg, 1986, p. 12) is suggested. Secondly, only the limited version of libertarian paternalism is justified, because the difference between nudging and manipulation is very slight. Although the conclusion from this book is that libertarian paternalism can be perceived as a consistent

approach, not everyone will agree with that. It is the result of differences in how freedom is defined. First and foremost, whether it is treated as an intrinsic value that cannot be compared to well-being.

The last few pages present synthetically the conclusions from the book. However, the reader may ask: what about the bold question that the author made at the beginning of the book? The question was: **‘Do the weaknesses of value-free economics imply the need for a paradigm shift?’** To begin with, it must be stated that the main question is purposefully not asked categorically (“imply the need”) because the purpose of this book is not to give the impression that we should completely dismiss value-free economics. Moreover, ‘paradigm shift’ is not perceived in this book as an overnight process.¹ From this perspective, value-free economics, despite its problems, will be used by many economists. The unwillingness to answer the question of dismissing value-free economics directly is also the result of the author’s awareness about a limited scope of the analysis presented in this book. The focus was on choices over time to point out to what problems the assumptions of revealed preference theory has led in methodology and in reality. From this perspective, value-free economics seems not to work properly and can be perceived as degenerative programme. However, it is only a partial analysis of value-free economics. Therefore, we cannot firmly state that value-free economics should be dismissed. Last but not least, not one single author but a scientific community decide about ‘paradigm shift.’ In that context, it seems that value-free economics is less and less prominent and the limits of this approach are known to most economists.

The most substantial criticism towards value-free economics presented in this book is that it has made some economists arrogant and dogmatic (due to the immunity from criticism).² This is why it seems that this paradigm is degenerative programme. Doing good science is about scrutiny and critical thinking (fallibilism). Therefore, we cannot have values or assumptions that cannot be doubted and this is just what economists have been looking for in the economic approach. Of course, criticism alone is never enough. In order to decide whether economists should stick to value-free economics or move towards value-based economics, the alternative conceptions to value-free economics were analysed.

The most attention was paid to normative approaches, because they are very critical towards value-free economics (especially towards revealed preference theory). Although they seem to resolve some issues in comparison to the revealed preference theory (e.g., measuring well-being, which is not connected

¹ The idea of Copernicus was not approved overnight; and it took some time for value-free economics to overcome institutional economics at the beginning of the 20th century

² Nowadays more and more economists criticize the arrogance of neoclassical economics and they argue that economics should be more humble, see (Colander, 2016; Hardt, 2021; Rodrik, 2015).

with choice), they have problems of their own. The most significant one is their fundamental relativity (e.g., establishing goals and what happiness means). However, I do hope that the analysis of the contemporary philosophy of science may make clear for a reader that it is impossible to escape from some relativity in science. In social sciences, the accusation of relativity is not a sufficient argument to dismiss normative approaches. Moreover, it seems that both assumptions: a) people do what is the best for them, b) there is one self over time, do not correspond with observable reality and they lead to many problems. Normative approaches have their challenges. However, it is better to have the unprecise answer to an important question (e.g., well-being, freedom) rather than not asking these questions at all (axiomatic revealed preference theory).

If normative approaches are better than value-free economics (both descriptively and predictively), it seems natural to ask why the ‘paradigm shift’ has not taken place and why normative approaches are not the new mainstream? When one analyses normative approaches, one can see how problematic they are and how many problems they generate because of their relativity. However, the purpose of this book was to demonstrate that value-free economics by using the axiomatic assumptions (e.g., utility, rationality) has not solved the problems investigated in normative approaches either. In the end, the contemporary philosophy of science highlights that one should be pragmatic instead of striving for the definitive answers and absolute Truth (e.g., (Feyerabend 1975; Lakatos, 1980; McCloskey, 1985; Rorty, 1979)). Therefore, scientists should choose the theories that are progressive and can resolve more problems. From the perspective of the topics examined in this book, normative approaches seem to have the potential for improvement, whereas value-free economics seems to be degenerative programme. Moreover, it appears that axiomatic value-free economics hinders the potential improvement of normative theories. It gives justification to some economists to treat normative approaches as fads or unscientific theories and to push them into the category of heterodox approaches (viz. insignificant).

The next reason why we have not observed ‘paradigm shift’ towards normative approaches is the complexity of these concepts. In science, simplicity is crucial (Mäki, 2009a; Mäki et al., 2017). The methodology of value-free economics is built on simplicity (a few axioms) and this is why, among other things, economic imperialism has been possible. At this moment normative theories indeed face many challenges and they are overcomplicated. However, with time and further work, some of these approaches can be improved and structured better. In this book, an attempt to synthesize the insights from normative approaches (multiple self conception, libertarian paternalism, capability approach, economics of happiness) has been undertaken. However, I am fully aware it is just a beginning, and there is still much to be done to achieve a satisfactory level of maturity of these normative concepts.

If I were pressed to answer 'yes' or 'no' to the question of 'paradigm shift', the answer would be 'yes.' The problems with value-free economics imply the need for 'paradigm shift,' but with many qualifications that were made along the way.

After answering the main question let us move to the implications of this book. Even though the part of this book analyses the history of economic thought, we need to have even more discussions on the foundations of value-free economics in the future. Otherwise, some scientists will still criticize this paradigm from descriptive perspective. The descriptive criticism of *homo economicus* (e.g., people are not rational and self-interested) misses the point when it comes to the real problems of value-free economics. What many opponents of value-free economics do it is the 'strife over methods' re-enact (positivism versus historicism). Economist community needs to know why value-free economics work as it works (methodology) to be able to move away from these unproductive discussions (*historia magistra vitae*).

A Weberian lens can be used to understand value-free economics as a science. Two concepts of Weber can be especially useful for economists: 1) ideal type, 2) value-freeness of science. The second concept is essential from the perspective of value-free economics. It is the case, because economists need to understand how to do science in the new scientific world which has been shaped by the contemporary philosophy of science. Economists must understand that the rules of the game have changed and that there are no objective facts or theories that are impregnated from the criticism (received view³).

If we cannot avoid using values in science, then we need to put them on the table and have the philosophical self-reflection on economics. To accomplish this, we can do a genealogy of value-free economics, and this book was partly about this process. But much more should be done. It is so important, because thanks to genealogy (historical analysis of the emergence of philosophical beliefs, see (Foucault, 1978; Nietzsche, 2006)), we can observe how economists have been trying to conceive values and their self-interests in the disguise of objective science. The genealogy of value-free economics can be a vaccination against the hubris of economists which is based on the belief in being positive science. It was indicated a few times through this book that the worst thing for scientists is to be uncritical in spirit and build theories that are completely immune from the criticism. By using axiomatic assumptions, value-free economists have tried to establish that their theories are not only true but they are objectively true.

A comprehensive understanding of the limits of economics should open the possibility to incorporate the insights from the other social sciences and

³ Position in the philosophy of science that identifies a scientific theory with a set of propositions which are considered to be linguistic objects, such as axioms. Received view is very similar to positivism and was very popular in economics, see (Hands, 2003; McCloskey, 1983).

humanities into economics (e.g., psychology, sociology, philosophy). It has been done over the last decades, but very often this knowledge was treated as a square in a round hole and it has not transformed economics. Economists have incorporated these insights but the ‘hard core’ of value-free economics have remained unchanged. It is worth studying, whether some paradigms from the other social sciences could change economics. The scrutiny of different paradigms is also important in the case of heterodox approaches which cannot gain enough appreciation because they struggle to have one simple and effective paradigm (Colander, & Su, 2018; Lawson, 2005).

The openness towards other sciences and different paradigms also means asking questions that neoclassical economists do not ask frequently. We should start the discussion on some fundamental assumptions in economics like autonomy and free will. Many economists do not want to have this discussion because these topics are treated as Pandora’s box. Asking the problematic questions can only open the discussion that cannot be resolved (as Robbins indicated in the case of ethics). However, these topics are a time bomb and not answering them will backfire in the future. Economists could see that not addressing some problematic topics does not make them go away (e.g., ethics, rationality, institutions, and methodology of economics).

Moreover, on the example of libertarian paternalism we observe that the lack of conversation on autonomy, its connection with free will and rationality results in many problems. Therefore, Sugden and Thaler talk past each other. They often cannot even agree to disagree. Scientists need to be explicit about their ‘methodological value judgments’ when they start the discussion. Otherwise, they will not understand each other.

The general goal of this book was to stir up an honest conversation on hard core assumptions that neoclassical economics which aspires to be value-free takes for granted (instrumental rationality, maximization of utility). Economists live in the turbulent times of paradigm shift, when it is known that the old paradigm of value-free economics meets numerous challenges and the increasing number of economists are looking for something new. These times imply the constant need for the philosophical self-reflection about the state of economics as a science. This book, as one among many works over the last years which criticizes value-free economics, has opened Pandora’s box to see what happens when we dismiss some axiomatic assumptions. It can be concluded that there are quite solid grounds for optimism. There can be a life after the value-free economics’ death.

References

- Adorno, T. (1973). *The jargon of authenticity* (K. Tarnowski, & F. Will, Trans.). Evanston, IL: Northwestern University Press.
- Agassi, J. (1988). *The gentle art of philosophical polemics*. La Salle, IL: Open Court.
- Ainslie, G. (1975). *Specious reward: A behavioral theory of impulsiveness and impulse control*. *Psychological Bulletin*, 82(4), 463–496.
- Ainslie, G. (2001). *Breakdown of will*. New York: Cambridge University Press.
- Akerlof, G. (1984). *An economic theorist's book of tales*. Cambridge: Cambridge University Press.
- Alexandrova, A. (2005). Subjective well-being and Kahneman's 'objective happiness'. *Journal of Happiness Studies*, 6(3), 301–324.
- Alexandrova, A. (2017). *A philosophy for the science of well-being*. Oxford: Oxford University Press.
- Alkire, S. (2005). Why the capability approach? *Journal of Human Development*, 6(1), 115–135.
- Alós-Ferrer, C., Strack, F. (2014). From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, (41), 1–11.
- Anderson, E. (2001). Symposium on Amartya Sen's philosophy: 2 Unstrapping the straitjacket of 'preference': A comment on Amartya Sen's contributions to philosophy and economics. *Economics & Philosophy*, 17(1), 21–38.
- Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, 102(7), 3333–3356.
- Andrews, F., & Robinson, J. (1991). Measures of subjective well-being. In J. Robinson, P. Shaver, & L. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 61–114). San Diego: Academic Press.
- Angner, E. (2010). Subjective well-being. *The Journal of Socio-Economics*, 39(3), 361–368.
- Angner, E. (2019). We're all behavioral economists now. *Journal of Economic Methodology*, 26(3), 195–207.
- Argyle, M. (1987). *The psychology of happiness*. London, New York: Methuen.
- Aristotle. (2000). *The Nicomachean ethics*. (R. Crisp, Trans. and Ed.). Cambridge: Cambridge University Press.
- Arrow, K. (1951). *Social choice and individual values*. New York: Wiley.
- Ashraf, N., Camerer, C., & Loewenstein, G. (2005). Adam Smith, behavioral economist. *Journal of Economic Perspectives*, 19(3), 131–145.
- Audi, R. (1990). Weakness of will and rational action. *Australasian Journal of Philosophy*, 68(3), 270–281.
- Backhouse, R. (1991). The neo-Walrasian research program in macroeconomics. In N. de Marchi, & M. Blaug (Eds.), *Appraising economic theories: Studies in the methodology of research programmes* (pp. 403–426). Cheltenham: Edward Elgar.

- Backhouse, R. (1994). The Lakatosian legacy in economic methodology. In R. Backhouse (Ed.), *New directions in economic methodology* (pp. 173–191). London: Routledge.
- Backhouse, R., & Cherrier, B. (2017). The age of the applied economist: The transformation of economics since the 1970s. *History of Political Economy*, (49), 1–33.
- Backhouse, R., Middleton, R., & Tribe, K. (1997). Economics is what economists do, but what do the numbers tell us? *Paper for Annual History of Economic Thought Conference*. Bristol: University of Bristol.
- Ballet J., Bazin, D., Dubois, J., & Mahieu, F. (2013). *Freedom, responsibility and economics of the person*. London: Routledge.
- Bargh, J., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462–479.
- Barrotta, P. (2008). Why economists should be unhappy with the economics of happiness. *Economics & Philosophy*, 24(2), 145–165.
- Baumeister, R. (1998). The self. In T. Daniel, T. Gilbert, & S. Fiske (Eds.), *Handbook of social psychology*, vol. 2 (pp. 680–740). New York: McGraw-Hill.
- Becker, G. (1960). An economic analysis of fertility. In Universities-National Bureau Committee for Economic Research, *Demographic and economic change in developed countries* (pp. 209–240). New York: Columbia University Press.
- Becker, G. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York: National Bureau of Economic Research.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217.
- Becker, G. (1974). A theory of social interactions. *Journal of Political Economy*, 82(6), 1063–93.
- Becker, G. (1976). *The economic approach to human behavior*. Chicago: University of Chicago Press.
- Becker, G. (1993). The economic way of looking at behavior. *The Journal of Political Economy*, 101(3), 385–409.
- Becker, G., & Murphy, K. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4), 675–700.
- Beed, C., & Beed, C. (2000). The status of economics as a naturalistic social science. *Cambridge Journal of Economics*, 24(4), 417–435.
- Beed, C., & Kane, O. (1991). What is the critique of the mathematization of economics? *Kyklos*, 44(4), 581–612.
- Benartzi, S., & Thaler, R. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives*, 21(3), 81–104.
- Bentham, J. (1907). *An introduction to the principles of morals and legislation*. Oxford: Clarendon Press. Retrieved May 26, 2019 from <https://oll.libertyfund.org/titles/278>
- Bentham, J. (1825). *The rationale of reward*. London: Published by John and H.L. Hunt. Retrieved May 20, 2019 from <https://archive.org/details/rationaleofrewar00bent/page/206>
- Bentham, J. (1983). *Deontology together with a table of the springs of action and article on utilitarianism*. (A. Goldworth, Ed.). Oxford: Clarendon Press.
- Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, 18(1), 133–165.

- Berlin, I. (2002). *Liberty: Incorporating "Four essays on liberty"* (2nd ed). Oxford: Oxford University Press.
- Bermúdez, J. (Ed.) (2018). *Self-control, decision theory, and rationality: New essays*. Cambridge: Cambridge University Press.
- Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2008). How are preferences revealed? *Journal of Public Economics*, 92(8-9), 1787–1794.
- Binder, M. (2014). Should evolutionary economists embrace libertarian paternalism? *Journal of Evolutionary Economics*, 24(3), 515–539.
- Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of Public Economics*, 88(7-8), 1359–1386.
- Blaug, M. (1975). Kuhn versus Lakatos, or paradigms versus research programmes in the history of economics. *History of Political Economy*, 7(4), 399–433.
- Blaug, M. (1978). *Economic Theory in Retrospect*. Cambridge: Cambridge University Press.
- Blaug, M. (1991). Afterword. In N. de Marchi, & M. Blaug (Eds.), *Appraising economic theories: Studies in the methodology of research programmes* (pp. 499–512). Cheltenham: Edward Elgar.
- Blaug, M. (1992). *The methodology of economics: Or, how economists explain* (2nd ed.). Cambridge: Cambridge University Press.
- Blaug, M. (1997). Ugly currents in modern economics. *Options Politiques*, 18(17), 3–8.
- Blaug, M. (2003). The formalist revolution of the 1950s. *Journal of the History of Economic Thought*, 25(2), 145–156.
- Boland, L. A. (1979). A critique of Friedman's critics. *Journal of Economic Literature*, 17(2), 503–522.
- Boland, L. A. (1981). On the futility of criticizing the neoclassical maximization hypothesis. *The American Economic Review*, 71(5), 1031–1036.
- Boland, L. A. (1997). *Critical economic methodology: A personal odyssey*. London: Routledge.
- Boldyrev, I., & Svetlova, E. (Eds.) (2016). *Enacting dismal science: New perspectives on the performativity of economics*. New York: Palgrave Macmillan.
- Boulding, K. E. (1969). Economics as a moral science. *The American Economic Review*, 59(1), 1–12.
- Braeutigam, S. (2005). Neuroeconomics – from neural systems to economic behaviour. *Brain Research Bulletin*, 67(5), 355–360.
- Brainerd, C., & Reyna, V. (2005). *The science of false memory*. New York: Oxford University Press.
- Brekke, K. A., Kverndokk, S., & Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics*, 87(9-10), 1967–1983.
- Brennan, T. J. (1989). A methodological assessment of multiple utility frameworks. *Economics & Philosophy*, 5(2), 189–208.
- Breton, E., & Sherlaw, W. (2011). Examining tobacco control strategies and aims through a social justice lens: An application of Sen's capability approach. *Public Health Ethics*, 4(2), 149–159.
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, 36(8), 917.
- Brisset, N. (2016). Economics is not always performative: Some limits for performativity. *Journal of Economic Methodology*, 23(2), 160–184.

- Brocas, I., & Carrillo, J. D. (2008). The brain as a hierarchical organization. *American Economic Review*, 98(4), 1312–1346.
- Broome, J. (1992). Deontology and economics. *Economics & Philosophy*, 8(2), 269–282.
- Bruni, L. (2008). Back to Aristotle? Happiness, eudaimonia and relational goods. In L. Bruni, F. Comim, & M. Pugno (Eds.). *Capabilities and happiness* (pp. 114–139). Oxford: Oxford University Press.
- Bruni, L., & Porta, P. L. (Eds.). (2005). *Economics and happiness: Framing the analysis*. Oxford: Oxford University Press.
- Bruni, L., & Porta, P. (Eds.). (2007). *Handbook on the economics of happiness*. Cheltenham: Edward Elgar.
- Bruni, L., & Sugden, R. (2007). The road not taken: How psychology was removed from economics, and how it might be brought back. *The Economic Journal*, 117(516), 146–173.
- Bruni, L., & Sugden, R. (2013). Reclaiming virtue ethics for economics. *Journal of Economic Perspectives*, 27(4), 141–164.
- Bruni, L., Comim, F., & Pugno, M. (Eds.). (2008). *Capabilities and happiness*. Oxford: Oxford University Press.
- Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, 2(1), 671–698.
- Buss, S. (1997). Weakness of will. *Pacific Philosophical Quarterly*, 78(1), 13–44.
- Butler, J. (1975). Of personal identity. In J. Perry (Ed.), *The analogy of religion* (pp. 99–105). Berkeley, CA: University of California Press.
- Caldwell, B. (1980). A critique of Friedman's methodological instrumentalism. *Southern Economic Journal*, 47(2), 366–374.
- Caldwell, B. (1982). *Beyond positivism: Economic methodology in the twentieth century*. London: Allen and Unwin.
- Caldwell, B. (1991). The methodology of scientific research programs in economics: Criticisms and conjectures. In G. Shaw (Ed.), *Economics, culture, and education: Essays in Honour of Mark Blaug* (pp. 95–107). Cheltenham: Edward Elgar.
- Callon, M. (2006). *What does it mean to say that economics is performative?* (CSI Working Papers Series, no. 005, Centre de Sociologie de l'Innovation (CSI), Mines ParisTech).
- Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review*, 151(3), 1211–1254.
- Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, 43(1), 9–64.
- Caplan, B. (2003). Stigler-Becker versus Myers-Briggs: Why preference-based explanations are scientifically meaningful and empirically important. *Journal of Economic Behavior & Organization*, 50(4), 391–405.
- Caplin, A., & Schotter, A. (Eds.) (2008). *The foundations of positive and normative economics: A handbook*. Oxford: Oxford University Press.
- Capps, J. (2019). The pragmatic theory of truth. *The Stanford Encyclopedia of Philosophy*. E. Zalta (Ed.). Retrieved May 5, 2019 from <https://plato.stanford.edu/archives/sum2019/entries/truth-pragmatic>

- Carlyle, T. (1843). *Past and present*. The Project Gutenberg. Retrieved May 5, 2019 from <https://www.gutenberg.org/ebooks/13534>
- Carlyle, T. (1850). *Latter-day pamphlets*. The Project Gutenberg. Retrieved May 5, 2019 from <http://www.gutenberg.org/ebooks/1140>
- Carnap, R. (1934). *The unity of science*. London: Kegan Paul, Trench, Trubner.
- Carnap, R. (1935). *Philosophy and logical syntax*. London: Kegan Paul, Trench, Trubner.
- Carnap, R. (1937). *Logical syntax of language*. London: Kegan Paul, Trench, Trubner.
- Carrera, M., Royer, H., Stehr, M., Sydnor, J., & Taubinsky, D. (2019). *How are preferences for commitment revealed?* (National Bureau of Economic Research Working Paper no. 26161). Retrieved August 3, 2019 from https://www.nber.org/system/files/working_papers/w26161/w26161.pdf
- Carter, I. (2018). Positive and negative liberty. *The Stanford Encyclopedia of Philosophy*, E. Zalta (Ed). Retrieved March 2, 2029 from <https://plato.stanford.edu/archives/sum2018/entries/liberty-positive-negative/>
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.
- Cedrini, M., & Fontana, M. (2018). Just another niche in the wall? How specialization is changing the face of mainstream economics. *Cambridge Journal of Economics*, 42(2), 427–451.
- Chmielewski, P. (2011). *Homo agens. Instytucjonalizm w naukach społecznych*. Warszawa: Poltext.
- Claassen, R. (2014). Capability paternalism. *Economics & Philosophy*, 30(1), 57–73.
- Clark, A. E., & Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, 104(424), 648–659.
- Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3), 359–381.
- Clark, J. M. (1918). Economics and modern psychology: I. *Journal of Political Economy*, 26(1), 1–30.
- Coddington, A. (1975). The rationale of general equilibrium theory. *Economic Inquiry*, 13(4), 539–558.
- Cohen, D. (2014). *Homo Economicus, the (lost) prophet of modern times*. Cambridge: Polity Press.
- Cohen, J. D., Ericson, K., Laibson, D., & White, J. M. (2020). Measuring time preferences. *Journal of Economic Literature*, 58(2), 299–347.
- Colander, D. (2000). The death of neoclassical economics. *Journal of the History of Economic Thought*, 22(2), 127–143.
- Colander, D. (2010). Moving beyond the rhetoric of pluralism. Suggestions for an 'inside-the-mainstream' heterodoxy. In R. F. Garnett Jr., E. K. Olsen, & M. Starr (Eds.), *Economic pluralism* (pp. 36–47). New York: Routledge.
- Colander, D. (2015). Why economics textbooks should, but don't, and won't change. *European Journal of Economics and Economic Policy: Intervention*, 12(2), 229–235.
- Colander, D. (2016). Creating humble economists. In G. DeMartino & D. McCloskey (Eds.). *The Oxford handbook of professional economic ethics* (pp. 737–749). Oxford: Oxford University Press.
- Colander, D., Holt, R. P., & Rosser, J. B. Jr. (2004). The changing face of mainstream economics. *Review of Political Economy*, 16(4), 485–499.
- Colander, D., Holt, R., & Rosser, J. B. (2011). The complexity era in economics. *Review of Political Economy*, 23(3), 357–369.

- Colander, D., & Su, H. C. (2015). Making sense of economists' positive-normative distinction. *Journal of Economic Methodology*, 22(2), 157–170.
- Colander, D., & Su, H. C. (2018). *How economics should be done: Essays on the art and craft of economics*. Cheltenham: Edward Elgar.
- Coleman, J. S. (1993). The impact of Gary Becker's work on sociology. *Acta Sociologica*, 36(3), 169–178.
- Collodel, M. (2016). Was Feyerabend a Popperian? Methodological issues in the history of the philosophy of science. *Studies in History and Philosophy of Science. Part A*, 57, 27–56.
- Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature*, 34(2), 669–700.
- Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, 54(5), 853–863.
- Cowell, F., & Witztum, A. (Eds.). (2007). *Lionel Robbins' essay on the nature and significance of economic science. 75th Anniversary Conference Proceedings*. London: Suntory and Toyota International Centres for Economics and Related Disciplines.
- Cowen, T. (1991). Self-constraint versus self-liberation. *Ethics*, 101(2), 360–373.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of Happiness Studies*, 4(2), 185–199.
- Czarny, B. (2010). *Pozytywizm a sądy wartościujące w ekonomii*. Monografie i Opracowania/Szkoła Główna Handlowa, no. 575, Warszawa: Oficyna Wydawnicza Szkoły Głównej Handlowej.
- Damasio, A. (2005). *Descartes' error: Emotion, reason, and the human brain*. New York: Penguin Books.
- Dasgupta, P. (2005). What do economists analyze and why: Values or facts? *Economics & Philosophy*, 21(2), 221–278.
- Dasgupta, P. (2008). Discounting climate change. *Journal of Risk and Uncertainty*, 37(2-3), 141–169.
- Daston, L., & Galison, P. (2007). *Objectivity*. New York: Zone Books.
- Davidson, D. (1980). How is weakness of the will possible? In D. Davidson, *Essays on actions and events* (pp. 21–42). Oxford: Clarendon Press.
- Davidson, R. (2003). Affective neuroscience and psychophysiology: Towards a synthesis. *Psychophysiology*, 40(5), 655–665.
- Davis, J. B. (1995). Personal identity and standard economic theory. *Journal of Economic Methodology*, 2(1), 35–52.
- Davis, J. B. (2003). *The theory of the individual in economics*. London: Routledge.
- Davis, J. B. (2004). *Identity and commitment: Sen's conception of the individual* (Tinbergen Institute Discussion Paper no. 04-055/2). Retrieved from <https://papers.tinbergen.nl/04055.pdf>
- Davis, J. B. (2005). Neoclassicism, artificial intelligence, and the marginalization of ethics. *International Journal of Social Economics*, 32(7), 590–601.
- Davis, J. B. (2006). The turn in economics: Neoclassical dominance to mainstream pluralism. *Journal of Institutional Economics*, 2(2), 1–20.
- Davis, J. B. (2008). The turn in recent economics and return of orthodoxy. *Journal of Economics*, 32(3), 349–366.

- Davis, J. B. (2010). *Individuals and identity in economics*. Cambridge: Cambridge University Press.
- Davis, J. B. (2016). Economics, neuroeconomics, and the problem of identity. *Schmollers Jahrbuch*, 136(1), 15–32.
- Davis, J., Smith, T., & Marsden, P. (2001). *General social survey: 1972–2000. Cumulative codebook*. Roper Center for Public Opinion Research. Retrieved May 2, 2019 from <https://ropercenter.cornell.edu/general-social-survey>
- De Beauvoir, S. (1949). *The ethics of ambiguity* (F. Bernard, Trans.). New York: Citadel Press.
- De Brigard, F. (2010). If you like it, does it matter if it's real? *Philosophical Psychology*, 23(1), 43–57.
- De Marchi, N., & Blaug, M. (Eds.). (1991). *Appraising economic theories: Studies in the methodology of research programs*. Cheltenham: Edward Elgar.
- De Sousa, R. (1990). *The rationality of emotion*. Cambridge, MA: The MIT Press.
- Debreu, G. (1991). The mathematization of economic theory. *The American Economic Review*, 81(1), 1–7.
- Decker, S., Elsner, W., & Flechtner, S. (Eds.). (2018). *Advancing pluralism in teaching economics: International perspectives on a textbook science*. London: Routledge.
- Dekker, E., & Remic, B. (2018). Two types of ecological rationality: Or how to best combine psychology and economics. *Journal of Economic Methodology*, 26(4), 291–306.
- Dennett, D. (1969). *Content and consciousness*. London: Routledge & Kegan Paul.
- Dennett, D. (1991). *Consciousness explained*. New York: Penguin Books.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542–575.
- Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. In E. Diener, & E. Suh (Eds.), *Culture and subjective well-being* (pp. 185–218). Cambridge, MA: The MIT Press.
- Diener, E., Scollon, C. N., & Lucas, R. E. (2004). The evolving concept of subjective well-being: The multifaceted nature of happiness. In P. T. Costa, & I. C. Siegler (Eds.), *Advances in cell aging and gerontology*, vol. 15 (pp. 187–220). Amsterdam: Elsevier.
- Dixon, W., & Wilson, D. (2006). Das Adam Smith Problem: A critical realist perspective. *Journal of Critical Realism*, 5(2), 251–272.
- Dixon, W., & Wilson, D. (2012). *A history of homo economicus: The nature of the moral in economics history*. London: Routledge.
- Dobusch, L., & Kapeller, J. (2009). “Why is economics not an evolutionary science?” New answers to Veblen’s old question. *Journal of Economic Issues*, 43(4), 867–898.
- Dobusch, L., & Kapeller, J. (2012). Heterodox united vs. mainstream city? Sketching a framework for interested pluralism in economics. *Journal of Economic Issues*, 46(4), 1035–1058.
- Donne, J. (1624). *No man is an island*. Retrieved April 7, 2018 from <https://web.cs.dal.ca/~johnston/poetry/island.html>
- Doris, J. M. (1998). Persons, situations, and virtue ethics. *Nous*, 32(4), 504–530.
- Dow, S. (2012). *Foundations for new economic thinking: A collection of essays*. London: Palgrave Macmillan.
- Dowell, R. S., Goldfarb, R. S., & Griffith, W. B. (1998). Economic man as a moral individual. *Economic Inquiry*, 36(4), 645–653.

- Drakopoulos, S. A., & Karayiannis, A. (2005). A review of Kuhnian and Lakatosian 'explanations' in economics. *History of Economic Ideas*, 13(2), 51–73.
- Dryden, J. (2018). Autonomy. *The Internet Encyclopedia of Philosophy*. Retrieved May 29, 2018 from <https://www.iep.utm.edu/autonomy/>
- Duflo, E. (2017). Richard T. Ely lecture: The economist as plumber. *American Economic Review*, 107(5), 1–26.
- Dworkin, G. (1972). Paternalism. *The Monist*, 56(1), 64–84.
- Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge: Cambridge University Press.
- Dzionek-Kozłowska, J. (2007). *System ekonomiczno-społeczny Alfreda Marshalla*. Warszawa: Wydawnictwo Naukowe PWN.
- Dzionek-Kozłowska, J. (2018). *Model homo oeconomicus. Geneza, ewolucja, wpływ na rzeczywistość gospodarczą*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. David, & M. Reder (Eds.), *Nations and households in economic growth: Essays in Honour of Moses Abramowitz* (pp. 89–125). New York: Academic Press.
- Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, 27(1), 35–47.
- Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The Economic Journal*, 111(473), 465–484.
- Edgeworth, F. Y. (1879). The hedonical calculus. *Mind*, 4, 394–408.
- Elias, N. (1991). *The society of individuals* (E. Jephcott, Trans.). Oxford: Basil Blackwell.
- Elsner, W. (2013). State and future of the 'citadel' and of the heterodoxies in economics: Challenges and dangers, convergences and cooperation. *European Journal of Economics and Economic Policies: Intervention*, 10(3), 286–298.
- Elster, J. (Ed.). (1987). *The multiple self*. Cambridge: Cambridge University Press.
- Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, 3(4), 99–117.
- Elster, J. (1996). Rationality and the emotions. *The Economic Journal*, 106(438), 1386–1397.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49(8), 709–724.
- Ericson, K. M., & Laibson, D. (2019). Intertemporal choice. In D. Bernheim, S. DellaVigna & D. Laibson (Eds.), *Handbook of behavioral economics: Applications and foundations 2* (pp. 1–67). Amsterdam: North-Holland.
- Etzioni, A. (1987). Toward a Kantian socio-economics. *Review of Social Economy*, 45(1), 37–47.
- Etzioni, A. (1988). *The moral dimension: Toward a new economics*. New York: The Free Press.
- Etzioni, A. (2011). Behavioral economics: Toward a new paradigm. *American Behavioral Scientist*, 55(8), 1099–1119.
- Fehr, E., Fischbacher, U., & Kosfeld, M. (2005). Neuroeconomic foundations of trust and social preferences: Initial evidence. *American Economic Review*, 95(2), 346–351.
- Feinberg, J. (1986). *Harm to self*. Oxford: Oxford University Press.

- Feyerabend, P. (1961). *Knowledge without foundations: Two lectures delivered on the Nellie Heldt lecture fund*. Oberlin, OH: Oberlin College.
- Feyerabend, P. (1993). *Against method* (3rd ed.). New York, London: Verso.
- Feyerabend, P. (2005). *Killing time: The autobiography of Paul Feyerabend*. Chicago: University of Chicago Press.
- Fiedor, B. (2010). Kryzys gospodarczy a kryzys ekonomii jako nauki. *Ekonomista*, (4), 453–466.
- Fiedor, B. (2013). Uwagi o potrzebie równowagi metodologicznej w ekonomii. *Studia Ekonomiczne*, 1(76), 101–118.
- Fiedor, B., & Ostapiuk, A. (2017). Utylitaryzm versus aksjologiczne i społeczne uwarunkowania wyborów ekonomicznych. In E. Mączyńska, & J. Sójka (Eds.). *Etyka i ekonomia – w stronę nowego paradygmatu* (pp. 19–44). Warszawa: Polskie Towarzystwo Ekonomiczne.
- Fine, B. (2019). Economics and interdisciplinarity: One step forward, n steps back? *Revista Crítica de Ciências Sociais*, (119), 131–148.
- Fleck, L. (2012). *Genesis and development of a scientific fact* (F. Bradley & T. J. Trenn, Trans.). Chicago: University of Chicago Press.
- Fletcher, G. (Ed.) (2015). *The Routledge handbook of philosophy of well-being*. Abingdon: Routledge.
- Fontana, M. (2014). Pluralism(s) in economics: Lessons from complexity and innovation. *Journal of Evolutionary Economics*, 24(1), 189–204.
- Fordyce, M. W. (1986). The *psychap inventory*: A multi-scale test to measure happiness and its concomitants. *Social Indicators Research*, 18(1), 1–33.
- Fordyce, M. W. (1988). A review of research on the happiness measures: A sixty second index of happiness and mental health. *Social Indicators Research*, 20(4), 355–381.
- Foster, J., & Sen, A. (1997) On economic inequality after a quarter of a century. In J. Foster, & A. Sen. (Eds.), *On economic inequality* (pp. 107–219). Oxford: Oxford University Press.
- Foucault, M. (1978). Nietzsche, genealogy, history. In J. Richardson, & B. Leiter (Eds.), *Nietzsche* (pp. 139–164). Oxford: Oxford University Press.
- Fourcade, M., Ollion, E., & Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1), 89–114.
- Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: WW Norton & Co.
- Frank, R. H. (1997). The frame of reference as a public good. *The Economic Journal*, 107(445), 1832–1847.
- Frank, R. H. (1999). *Luxury fever: Why money fails to satisfy in an era of excess*. New York: The Free Press.
- Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20.
- Frankfurt, H. (1988). *The importance of what we care about*. Cambridge: Cambridge University Press.
- Frederick, S., & Loewenstein, G. (1999). Hedonic adaptation. In D. Kahneman, E. Diener, & N. Schwarz. (Eds.). *Well-being: The foundations of hedonic psychology* (pp. 302–329). New York: Russell Sage Foundation.

- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40(2), 351–401.
- Freud, S. (1956a). Formulations on the two principles of mental functioning. In J. Strachey, & A. Freud (Eds.), *The standard edition of the complete psychological works of Sigmund Freud*, vol. 12. London: Hogarth.
- Freud, S. (1956b). The Ego and the Id. In J. Strachey, & A. Freud (Eds.), *The standard edition of the complete psychological works of Sigmund Freud*, vol. 19, London: Hogarth.
- Frey, B. S. (2010). *Happiness: A revolution in economics*. Cambridge, MA: The MIT Press.
- Frey, B. S. (2018). *Economics of happiness*. New York: Springer.
- Frey, B. S., Benz, M., & Stutzer, A. (2004). Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics JITE*, 160(3), 377–401.
- Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5), 589–611.
- Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, 87(4), 746–755.
- Frey, B. S., & Stutzer, A. (2002a). *Happiness and economics: How the economy and institutions affect human well-being*. Princeton: Princeton University Press.
- Frey, B. S., & Stutzer, A. (2002b). What can economists learn from happiness research? *Journal of Economic literature*, 40(2), 402–435.
- Frey, B. S., & Stutzer, A. (2010). Happiness and public choice. *Public Choice*, 144(3–4), 557–573.
- Friedman, M. (1953). *Essays in positive economics*. Chicago: University of Chicago Press.
- Friedman, M. (1962). *Capitalism and freedom*. Chicago: University of Chicago Press.
- Friedman, M., & Friedman, R. (1980). *Free to choose: A personal statement*. Boston: Houghton Mifflin Harcourt.
- Fromm, E. (1994). *Escape from freedom*. New York: Henry Holt & Co.
- Fuchs, V. R. (1994). Nobel laureate: Gary S. Becker. Ideas about facts. *Journal of Economic Perspectives*, 8(2), 183–192.
- Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review*, 96(5), 1449–1476.
- Fudenberg, D., & Levine, D. K. (2012). Timing and self-control. *Econometrica*, 80(1), 1–42.
- Fumagalli, R. (2016). Decision sciences and the new case for paternalism: Three welfare-related justificatory challenges. *Social Choice and Welfare*, 47(2), 459–480.
- Galbraith, J. (1998). *The affluent society*. Boston: Mariner Books.
- Garrett, B. (1998). *Personal identity and self-consciousness*. London: Routledge.
- Gellner, E. (1960). *Words and things*. Boston: Beacon Press.
- Gigerenzer, G. (2015). On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology*, 6(3), 361–383.
- Gilbert, D. (2007). *Stumbling on happiness*. Vintage.
- Gill, N., & Gill, M. (2012). The limits to libertarian paternalism: Two new critiques and seven best-practice imperatives. *Environment and Planning C: Government and Policy*, 30(5), 924–940.

- Glaeser, E. (2006). Paternalism and psychology. *University of Chicago Law Review*, 73(1), 133–156.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29(1), 1–17.
- Godłów-Legiędź, J. (2003). Nowa ekonomia instytucjonalna: nowe spojrzenie na istotę gospodarowania i rozwój. *Acta Universitatis Lodzianis. Folia Oeconomica*, (169), 59–78.
- Godłów-Legiędź, J. (2008). Kontrowersje wokół pomiaru dobrobytu społecznego. *Ruch Prawniczy, Ekonomiczny i Socjologiczny*, (4), 91–108.
- Godłów-Legiędź, J. (2010). *Współczesna ekonomia: ku nowemu paradygmatowi*. Warszawa: Wydawnictwo C.H. Beck.
- Godłów-Legiędź, J. (2014). Ekonomia i inżynieria społeczna z perspektywy kryzysu. *Ekonomista*, (1), 9–26.
- Goethe, J. W. (1961). *Faust* (W. Kaufmann, Trans.) New York: Anchor Books.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Garden City.
- Goldfarb, R., & Griffith, W. (1991a). Amending the economist's 'rational egoist' model to include moral values and norms. Part 1: The problem. In K. Koford, & J. Miller (Eds.), *Social norms and economic institutions* (pp. 39–57). Ann Arbor, MI: University of Michigan Press.
- Goldfarb, R., & Griffith, W. (1991b). Amending the economist's 'rational egoist' model to include moral values and norms. Part 2: Alternative solutions. In K. Koford, & J. Miller (Eds.), *Social Norms and Economic Institutions* (pp. 59–84). Ann Arbor, MI: University of Michigan Press.
- Gonzalez, W. (2014). The evolution of Lakatos's repercussion on the methodology of economics. *HOPOS: The Journal of the International Society for the History of Philosophy of Science*, 4(1), 1–25.
- Goźdź, M., Hardt, Ł., & Kwarciański, T. (Eds.). (2016). *Metaekonomia: zagadnienia z filozofii ekonomii*. Kraków: Copernicus Center Press.
- Graham, C. (2009). *Happiness around the world: The paradox of happy peasants and miserable millionaires*. Oxford: Oxford University Press.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510.
- Graupe, S. (2019). "Waging the war of ideas": Economics as a textbook science and its possible influence on human minds. In S. Decker, W. Elsner, S. Flechtner (Eds.), *Advancing pluralism in teaching economics* (pp. 173–191). London: Routledge.
- Griffin, J. (1986). *Well-being: Its meaning, measurements, and moral importance*. Oxford: Oxford University Press.
- Grüne-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, 38(4), 635–645.
- Gul, F., & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, 69(6), 1403–1435.
- Gul, F., & Pesendorfer, W. (2004). Self-control and the theory of consumption. *Econometrica*, 72(1), 119–158.
- Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin, & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 3–42). Oxford: Oxford University Press.
- Hagerty, M. R., & Veenhoven, R. (2003). Wealth and happiness revisited – growing national income does go with greater happiness. *Social Indicators Research*, 64(1), 1–27.

- Hammersley, M. (2017). On the role of values in social research: Weber vindicated? *Sociological Research Online*, 22(1), 1–12.
- Hands, D. W. (1984). The role of crucial counterexamples in the growth of economic knowledge: Two case studies in the recent history of economic thought. *History of Political Economy*, 16(1), 59–67.
- Hands, D. W. (1990). Second thoughts on 'second thoughts': Reconsidering the Lakatosian progress of the general theory. *Review of Political Economy*, 2(1), 69–81.
- Hands, D. W. (1991). The problem of excess content: Economics, novelty and a long Popperian tale. In N. de Marchi, & M. Blaug (Eds.), *Appraising economic theories: Studies in the methodology of research programmes* (pp. 58–75). Cheltenham: Edward Elgar.
- Hands, D. W. (1993). Popper and Lakatos in economic methodology. In B. Gustafsson, C. Knudsen, & U. Mäki (Eds.), *Rationality, institutions and economic methodology* (pp. 61–75). London: Routledge.
- Hands, D. W. (2001). *Reflection without rules: Economic methodology and contemporary science theory*. Cambridge: Cambridge University Press.
- Hands, D. W. (2003). Reconsidering the received view of the 'Received View': Kant, Kuhn, and the demise of positivist philosophy of science. *Social Epistemology*, 17(2-3), 169–173.
- Hands, D. W. (2007). Effective tension in Robbins's economic methodology. In F. Cowell, & A. Witztum (Eds.) *Lionel Robbins' essay on the nature and significance of economic science. 75th anniversary conference proceedings* (pp. 152–168). London: London School of Economics and Political Science.
- Hands, D. W. (2009). Economics, psychology and the history of consumer choice theory. *Cambridge Journal of Economics*, 34(4), 633–648.
- Hands, D. W. (2012). The positive-normative dichotomy and economics. In U. Mäki (Ed.), *Philosophy of economics* (pp. 219–39). Amsterdam: Elsevier.
- Hands, D. W. (2014). Paul Samuelson and revealed preference theory. *History of Political Economy*, 46(1), 85–116.
- Harding, S. (Ed.). (1975). *Can theories be refuted? Essays on the Duhem-Quine thesis*. Springer.
- Hardt, Ł. (2013). *Studia z realistycznej filozofii ekonomii*. Warszawa: Wydawnictwo C.H. Beck.
- Hardt, Ł. (2016). Modele w ekonomii. In M. Gorazda, Ł. Hardt, & T. Kwarciński (Eds.), *Metaekonomia. Zagadnienia z filozofii ekonomii* (pp. 223–250). Kraków: Copernicus Center Press.
- Hardt, Ł. (2017). *Economics without laws: Towards a new philosophy of economics*. London: Palgrave Macmillan.
- Hardt, Ł. (2020). Utylitaryzm, deontologia i etyka cnót: zbieżne czy przeciwstawne fundamenty etyczne ekonomii? *Ekonomista*, (2), 249–265.
- Hardt Ł. (2021). An essay on humble economics. In P. Róna, L. Zsolnai, & A. Wincewicz-Price (Eds.), *Words, objects and events in economics: The making of economic theory* (pp. 13–32). Springer. Retrieved May 2, 2020 from <https://www.springer.com/gp/book/9783030526726>
- Hare, R. (1952). *The language of morals*. Oxford: Clarendon Press.
- Hare, R. (1963). *Freedom and reason*. Oxford: Clarendon Press.
- Harris, S. (2011). *The moral landscape: How science can determine human values*. New York: Simon and Schuster.
- Harris, S. (2012). *Free will*. New York: The Free Press.

- Harrison, G. W. (2008). Neuroeconomics: A critical reconsideration. *Economics & Philosophy*, 24(3), 303–344.
- Harsanyi, J. C. (1977). Rule utilitarianism and decision theory. *Erkenntnis*, 11(1), 25–53.
- Hausman, D. M. (1992). *The inexact and separate science of economics*. Cambridge: Cambridge University Press.
- Hausman, D. M. (1998). Problems with realism in economics. *Economics & Philosophy*, 14(2), 185–213.
- Hausman, D. M. (Ed.). (2007). *The philosophy of economics: An anthology* (3rd ed.). Cambridge: Cambridge University Press.
- Hausman, D. M. (2011). *Why satisfy preferences?* (Papers on Economics and Evolution no. 2011-24).
- Hausman, D. M. (2012). *Preference, value, choice, and welfare*. Cambridge, New York: Cambridge University Press.
- Hausman, D. M. (2016). On the econ within. *Journal of Economic Methodology*, 23(1), 26–32.
- Hausman, D. M., & McPherson, M. (2006). *Economic analysis, moral philosophy and public policy*. Cambridge: Cambridge University Press.
- Hausman, D. M., McPherson, M., & Satz, D. (2017). *Economic analysis, moral philosophy, and public policy*. Cambridge: Cambridge University Press.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1), 123–136.
- Hayek, F. (1978). *The constitution of liberty*. Chicago: The University of Chicago Press.
- Hayek, F. (1988). *The fatal conceit*. Chicago: University of Chicago Press.
- Hechter, M. (1992). Should values be written out of the social scientist's lexicon? *Sociological Theory*, 10(2), 214–230.
- Hechter, M. (1994). The role of values in rational choice theory. *Rationality and Society*, 6(3), 318–333.
- Hédoin, C. (2013). *Sen on rationality, commitment and preferences* (REGARDS – University of Reims Champagne-Ardenne Working Paper).
- Hédoin, C. (2017). Normative economics and paternalism: The problem with the preference-satisfaction account of welfare. *Constitutional Political Economy*, 28(3), 286–310.
- Hegel, G. (2002). *The phenomenology of mind*. London: Routledge.
- Heidegger, M. (1962). *Being and time* (J. Macquarrie, E. Robinson, Trans.). New York: Harper & Row.
- Heilmann, C. (2010). *Rationality and time: A multiple-self model of personal identity over time for decision and game theory* (Doctoral dissertation). London: London School of Economics and Political Science. Retrieved May 9, 2019 from <http://etheses.lse.ac.uk/2210/1/U613442.pdf>
- Hempel, C. G. (1945). Studies in the logic of confirmation (I). *Mind*, 54(213), 1–26.
- Hesse, H. (2000). *The glass bead game*. New York: Random House.
- Hetherington, S. (n.d.). Fallibilism. *The Internet Encyclopaedia of Philosophy*. Retrieved April 13, 2019 from <https://www.iep.utm.edu/fallibil/>
- Hobbes, T. (1655). *The elements of philosophy: De corpore*. Retrieved May 9, 2018 from <https://archive.org/details/englishworksth01hobbgoog/page/n9>

- Hoch, S. J., & Loewenstein, G. F. (1991). Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research*, 17(4), 492–507.
- Hodgson, G. M. (1998). The approach of institutional economics. *Journal of Economic Literature*, 36(1), 166–192.
- Hodgson, G. M. (2000). What is the essence of institutional economics? *Journal of Economic Issues*, 34(2), 317–329.
- Hodgson, G. M. (2007). Meanings of methodological individualism. *Journal of Economic Methodology*, 14(2), 211–226.
- Hodgson, G. M. (2012a). *From pleasure machines to moral communities: An evolutionary economics without homo economicus*. Chicago: University of Chicago Press.
- Hodgson, G. M. (2012b). On the limits of rational choice theory. *Economic Thought*, 1(1), 94–108.
- Hollis, M. (1983). Rational preferences. *The Philosophical Forum*, 14(3), 246–262.
- Holton, R. (1999). Intention and weakness of will. *The Journal of Philosophy*, 96(5), 241–262.
- Holton, R. (2003). How is strength of will possible. In T. Stroud, & C. Tappolet (Eds.), *Weakness of will and practical irrationality* (pp. 39–67). Oxford: Clarendon Press.
- Hoyningen-Huene, P. (2017). *Revisiting Friedman's F53. Popper, Knight, and Weber*. Retrieved August 29, 2019 from <http://philsci-archive.pitt.edu/12910/1/Friedmans%20F53.pdf>
- Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, 133(1), 23–30.
- Hudson, H. (2001). *A materialist metaphysics of the human person*. Ithaca, NY: Cornell University Press.
- Hume, D. (1896). *A treatise of human nature*. Oxford: Clarendon Press.
- Hutchison, T. (1938). *The significance and basic postulates of economic theory*. London: Macmillan.
- Hutchison, T. (1981). *The politics and philosophy of economics*. New York: New York University Press.
- Hutchison, T. (1996). On the relations between philosophy and economics. Part I: Frontier problems in an era of departmentalized and internationalized 'professionalism'. *Journal of Economic Methodology*, 3(2), 187–213.
- Huxley, A. (2006). *Brave new world*. New York: Harper Perennial.
- Infante, G., Lecouteux, G., & Sugden, R. (2016). Preference purification and the inner rational agent: A critique of the conventional wisdom of behavioural welfare economics. *Journal of Economic Methodology*, 23(1), 1–25.
- Inglehart, R., Basanez, M., Diez-Medrano, J., Halman, L., & Luijckx, R. (2000). *World values surveys and European values surveys, 1981–1984, 1990–1993, and 1995–1997* (ICPSR 2790). Ann Arbor, MI: Institute for Social Research. Retrieved May 5, 2019 from https://web.stanford.edu/group/ssds/dewidocs/icpsr2790_superseded/cb2790.pdf
- James, W. S. (1981). *Pragmatism: A new name for some old ways of thinking*. Indianapolis, IN: Hackett Publishing.
- Jeffrey, R. (1974). Preference among preferences. *The Journal of Philosophy*, 71(13), 377–391.
- Jevons, W. S. (1911). *Political economy*. London: Macmillan.
- Johnston, M. (1987). Human beings. *The Journal of Philosophy*, 84(2), 59–83.

- Juster, T. (1985). Preference for work and leisure. In T. Juster, & F. Stafford (Eds.), *Time, goods, and well-being* (pp. 333–351). Ann Arbor, MI: University of Michigan.
- Kahneman, D. (2003a). A psychological perspective on economics. *The American Economic Review*, 93(2), 162–168.
- Kahneman, D. (2003b). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93(5), 1449–1475.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
- Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: Foundations of hedonic psychology*. New York: Russell Sage Foundation.
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 76(4), 728–741.
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702), 1776–1780.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2), 375–406.
- Kalberg, S. (Ed.) (2005). *Max Weber: Readings and commentary on modernity*. New York, Oxford: Wiley-Blackwell.
- Kaldor, N. (1957). A model of economic growth. *The Economic Journal*, 67(268), 591–624.
- Kant, I. (1993). *Grounding for the metaphysics of morals*. (J. W. Ellington, Transl.). Indianapolis/Cambridge: Hackett Publishing.
- Kant, I. (1996). *The Metaphysics of Morals* (M. Gregor, Trans.). Cambridge: Cambridge University Press.
- Kennett, J. (2001). *Agency and responsibility: A common-sense moral psychology*. Oxford: Clarendon Press.
- Kenny, C. (1999). Does growth cause happiness, or does happiness cause growth? *Kyklos*, 52(1), 3–25.
- Kesebir, P., & Diener, E. (2008). In defense of happiness: Why policymakers should care about subjective well-being. In L. Bruni, F. Comim, & M. Pugno (Eds.). *Capabilities and happiness* (pp. 69–90). Oxford: Oxford University Press.
- Keynes, J. (1999). *The scope and method of political economy*. Kitchener, Ontario: Batoche.
- Kierkegaard, S. (1980). *The sickness unto death: The Christian psychological exposition for upbuilding and awakening* (H. Hong, & E. Hong, Trans.). Princeton: Princeton University Press.
- Kierkegaard, S. (1983). *Fear and trembling and repetition*. Princeton: Princeton University Press.
- Kierkegaard, S. (1992). *Concluding unscientific postscript to "Philosophical fragments"*. Princeton: Princeton University Press.
- Kincaid, H., Dupré, J., & Wylie, A. (Eds.). (2007). *Value-free science: Ideals and illusions?* Oxford: Oxford University Press.

- Kincaid, H., & Ross, D. (Eds.). (2009). *The Oxford handbook of philosophy of economics*. Oxford Handbooks. Oxford: Oxford University Press.
- Kirchgässner, G. (2013). The weak rationality principle in economics. *Swiss Journal of Economics and Statistics*, 149(1), 1–26.
- Klappholz, K. (1964). Value judgments and economics. *The British Journal for the Philosophy of Science*, 15(58), 97–114.
- Klappholz, K. (1991). Comment on Shearmur. In N. de Marchi, & M. Blaug (Eds.), *Appraising economic theories: Studies in the methodology of research programmes* (pp. 53–57). Cheltenham: Edward Elgar.
- Kleist, C. (2018). Global ethics: Capabilities approach. *The Internet Encyclopaedia of Philosophy*. Retrieved May 29, 2018 from <http://www.iep.utm.edu/ge-capab/#H3/>
- Klimczak, B. (2014). Aksjologiczne uwikłanie ekonomii. *Annales. Etyka w Życiu Gospodarczym*, 17(1), 9–21.
- Knight, F. (1921). *Risk, uncertainty and profit*. Boston, New York: Houghton Mifflin.
- Kolak, D., & Martin, R. (1991). *Self and identity: Contemporary philosophical issues*. New York: Macmillan.
- Korfmacher, C. (2018). Personal identity. *The Internet Encyclopedia of Philosophy*. Retrieved May 30, 2028 from <http://www.iep.utm.edu/person-i/>
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.
- Kwarciański, T. (2016). Zagadnienia etyczne w ekonomii. In M. Gorazda, Ł. Hardt, & T. Kwarciański (Eds.), *Metaekonomia. Zagadnienia z filozofii ekonomii* (pp. 361–386). Kraków: Copernicus Center Press.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2), 443–477.
- Laibson, D. (2015). Why don't present-biased agents make commitments? *American Economic Review*, 105(5), 267–72.
- Lakatos, I. (1980). *The methodology of scientific research programmes*. Philosophical Papers, Volume 1. Cambridge: Cambridge University Press.
- Lawson, T. (1997). *Economics & reality*. Abingdon: Routledge.
- Lawson, T. (2003). *Reorienting economics*. London, New York: Routledge.
- Lawson, T. (2005). The nature of heterodox economics. *Cambridge Journal of Economics*, 30(4), 483–505.
- Lawson, T. (2015). *Essays on: The nature and state of modern economics*. Abingdon: Routledge.
- Layard, R. (2003). *Happiness: Has social science a clue?* Lionel Robbins Memorial Lectures 2002/3. London: London School of Economics.
- Layard, R. (2005). *Happiness: Lessons from a new science*. London: Penguin Books.
- Lazear, E. P. (2000). Economic imperialism. *The Quarterly Journal of Economics*, 115(1), 99–146.
- Leibenstein, H. (1966). Allocative efficiency vs. "X-efficiency". *The American Economic Review*, 56(3), 392–415.
- Leibenstein, H. (1979). A branch of economics is missing: Micro-micro theory. *Journal of Economic Literature*, 17(2), 477–502.

- Leijonhufvud, A. (1973). Life among the Econ. *Economic Inquiry*, 11(3), 327–337.
- Lepper, H. S. (1998). Use of other-reports to validate subjective well-being measures. *Social Indicators Research*, 44(3), 367–379.
- Lepper, M. R., & Greene, D. (Eds.). (1978). *The hidden costs of reward*. Hillsdale, NJ: Erlbaum.
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York: Harper Collins.
- Lewis, D. (1976). Survival and identity. In A. Rorty (Ed.), *The identities of persons* (pp. 17–40). Berkeley, CA: University of California Press.
- Lieberman, M. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. Forgas, K. Williams, & W. von Hippel (Eds.), *Social judgments: Explicit and implicit processes* (pp. 44–67). New York: Cambridge University Press.
- Lipsey, R. (2001). Successes and failures in the transformation of economics. *Journal of Economic Methodology*, 8(2), 169–201.
- Locke, J. (1689). *An essay concerning human understanding*. Retrieved May 7, 2018 from <https://oll.libertyfund.org/titles/761>
- Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65(3), 272–292.
- Loewenstein, G., & Elster, J. (Eds.). (1992). *Choice over time*. New York: Russell Sage Foundation.
- Loewenstein, G., & O'Donoghue, T. (2004). *Animal spirits: Affective and deliberative processes in economic behavior* (CAE Working Paper no. 04-14, Carnegie Mellon). Retrieved March 3, 2018 from <https://cae.economics.cornell.edu/04-14.pdf>
- Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics*, 118(4), 1209–1248.
- Loewenstein, G., & Schkade, D. (1999). Wouldn't it be nice? Predicting future feelings. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 81–105). New York: Russell Sage Foundation.
- Loewenstein, G., & Ubel, P. A. (2008). Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics*, 92(8-9), 1795–1810.
- Lutz, M. A. (1993). The utility of multiple utility: A comment on Brennan. *Economics & Philosophy*, 9(1), 145–154.
- Machlup, F. (1978). *Methodology of economics and other social sciences*. New York: Academic Press.
- MacIntyre, A. (1996). *A short history of ethics*. New York: Touchstone.
- MacKenzie, D. (2006). Is economics performative? Option theory and the construction of derivatives markets. *Journal of the History of Economic Thought*, 28(1), 29–55.
- MacKenzie, D. A., Muniesa, F., & Siu, L. (Eds.). (2007). *Do economists make markets? On the performativity of economics*. Princeton: Princeton University Press.
- Madra, Y. (2016). *Late Neoclassical Economics: The restoration of theoretical humanism in contemporary economic theory*. Abingdon: Routledge.
- Mäki, U. (2002). *Fact and fiction in economics: Models, realism and social construction*. Cambridge: Cambridge University Press.

- Mäki, U. (2006). On the method of isolation in economics. In J. Davis (Ed.), *Recent developments in economic methodology* (pp. 3–37). Cheltenham: Edward Edgar.
- Mäki, U. (2009a). Economics imperialism: Concept and constraints. *Philosophy of the Social Sciences*, 39(3), 351–380.
- Mäki, U. (2009b). Unrealistic assumptions and unnecessary confusions: Rereading and rewriting F53 as a realist statement. In U. Mäki (Ed.), *The methodology of positive economics. Reflections on the Milton Friedman legacy* (pp. 90–116). Cambridge: Cambridge University Press.
- Mäki, U., Walsh, A., & Pinto, M. F. (Eds.). (2017). *Scientific imperialism: Exploring the boundaries of interdisciplinarity*. London: Routledge.
- Marchionni, C., & Vromen, J. (Eds.). (2014). *Neuroeconomics: Hype or hope?* Abingdon: Routledge.
- Marshall, A. (1920). *Principles of economics*. Library of Economics and Liberty. Retrieved May 3, 2017 from <http://www.econlib.org/library/Marshall/marP3.html>
- Martin, R. (1987). *The meaning of language*. Cambridge, MA: The MIT Press.
- Martin, R., & Barresi, J. (Eds.). (2003). *Personal identity*. Hoboken, NJ: Blackwell Publishing.
- Marx, K. (2006). *Wage-labour and capital & Value, price and profit*. New York: International Publishers.
- McCloskey, D. N. (1983). The rhetoric of economics. *Journal of Economic Literature*, 21(2), 481–517.
- McCloskey, D. N. (1985). *The rhetoric of economics*. Madison, WI: University of Wisconsin Press.
- McCloskey, D. N. (2008). Adam Smith, the last of the former virtue ethicists. *History of Political Economy*, 40(1), 43–71.
- McIntyre, A. (1990). Is akratic action always irrational? In O. Flanagan, A Rorty (Eds.), *Identity, character, and morality* (pp. 379–400). Cambridge, MA: The MIT Press.
- McQuillin, B., & Sugden, R. (2012). How the market responds to dynamically inconsistent preferences. *Social Choice and Welfare*, 38(4), 617–634.
- Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology*, 69(4), 603–610.
- Mele, A. R. (1987). *Irrationality: An essay on akrasia, self-deception and self-control*. Oxford: Oxford University Press.
- Mele, A. R. (2002). Akratics and addicts. *American Philosophical Quarterly*, 39(2), 153–167.
- Mele, A. R. (2010). Weakness of will and akrasia. *Philosophical Studies*, 150(3), 391–404.
- Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, 126(6), 910–924.
- Metcalf, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological Review*, 106(1), 3–19.
- Michalos, A. C. (1991). *Global report on student well-being*. Vol. 1: *Life satisfaction and happiness*. New York: Springer.
- Michoń, P. (2010). *Ekonomia szczęścia*. Poznań: Dom Wydawniczy Harasimowicz.
- Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378.

- Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn. Retrieved from <https://archive.org/details/a592840000milluoft>
- Mill J. S. (1874). On the definition of political economy; and on the method of investigation proper to it. In *Essays on some unsettled questions of political economy* (2nd ed.). London: Longmans, Green, Reader & Dyer.
- Mill, J. S. (1909). *Autobiography*. The Harvard Classics, vol. 25 (C.E. Norton, Ed.). New York: P.F. Collier & Son.
- Mill, J. S. (2003). 'Utilitarianism' and 'On liberty'. Including 'Essay on Bentham' and selections from the writings of Jeremy Bentham and John Austin (M. Warnock, Ed.). Malden, MA: Blackwell.
- Miller, D. (1974). Popper's qualitative theory of verisimilitude. *The British Journal for the Philosophy of Science*, 25(2), 166–177.
- Minkler, L. (1999). The problem with utility: Toward a non-consequentialist/utility theory synthesis. *Review of Social Economy*, 57(1), 4–24.
- Mirowski, P. (1991). *More heat than light: Economics as social physics, physics as nature's economics*. Cambridge: Cambridge University Press.
- Mirowski, P. (2002). *Machine dreams: Economics becomes a cyborg science*. Cambridge: Cambridge University Press.
- Mirowski, P. (2004). The scientific dimensions of social knowledge and their distant echoes in 20th-century American philosophy of science. *Studies in History and Philosophy of Science Part A*, 35(2), 283–326.
- Mitchell, G. (2005). Libertarian paternalism is an oxymoron. *Northwestern University Law Review*, 99(3), 1248–1276.
- Modigliani, F. (1966). The life cycle hypothesis of saving, the demand for wealth and the supply of capital. *Social Research*, 33(2), 160–217.
- Mongin, P. (2001). Value judgments and value neutrality in economics: A perspective from today. Retrieved May 5, 2019, from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.2887&rep=rep1&type=pdf>
- Mongin, P. (2006). Value judgments and value neutrality in economics. *Economica*, 73(290), 257–286.
- Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about people's folk concept of free will. *Review of Philosophy and Psychology*, 1(2), 211–224.
- Morgan, M. S. (2006). Economic man as model man: Ideal types, idealization and caricatures. *Journal of the History of Economic Thought*, 28(1), 1–27.
- Morgan, M. S. (2012). *The world in the model: How economists work and think*. Cambridge, New York: Cambridge University Press.
- Myers, D. G. (1992). *The pursuit of happiness*. New York: William Morrow and Company.
- Myrdal, G. (1970). *Objectivity in social research*. London: Gerald Duckworth.
- Nagel, E. (1961). *The structure of science. Problems in the logic of scientific explanation*. London: Routledge & Kegan Paul.
- Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.
- Nettle, D. (2005). *Happiness: The science behind your smile*. Oxford: Oxford University Press.

- Ng, Y. K. (1997). A case for happiness, cardinalism, and interpersonal comparability. *The Economic Journal*, 107(445), 1848–1858.
- Nichols, S., & Bruno, M. (2010). Intuitions about personal identity: An empirical study. *Philosophical Psychology*, 23(3), 293–312.
- Nietzsche, F. (2006). *On the Genealogy of Morality and other writings: Revised student edition* (2nd ed.). Cambridge: Cambridge University Press.
- Noonan, H. W. (2004). *Personal identity*. London: Routledge.
- North, D. C. (1978). Structure and performance: The task of economic history. *Journal of Economic Literature*, 16(3), 963–978.
- Nozick, R. (1974). *Anarchy, state, and utopia*. New York: Basic Books.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge, MA: Harvard University Press.
- Nussbaum, M. C. (1992). Human functioning and social justice: In defense of Aristotelian essentialism. *Political Theory*, 20(2), 202–246.
- Nussbaum, M. C. (1995). Human capabilities, female human beings. In M. Nussbaum, & J. Glover (Eds.), *Women, culture, and development: A study of human capabilities*. Oxford: Clarendon Press.
- Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*. Cambridge: Cambridge University Press.
- Nussbaum, M. C. (2005). Mill between Aristotle and Bentham. In L. Bruni, & P. Porta (Eds.), *Economics and happiness* (pp. 170–183). Oxford: Oxford University Press.
- Nussbaum, M. C. (2011). *Creating capabilities*. Cambridge, MA: Harvard University Press.
- Oddie, G. (1986). The poverty of the Popperian program for truthlikeness. *Philosophy of Science*, 53(2), 163–178.
- O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1), 103–124.
- O'Donoghue, T., & Rabin, M. (2001). Choice and procrastination. *The Quarterly Journal of Economics*, 116(1), 121–160.
- Offer, A. (2006). *The challenge of affluence: Self-control and well-being in the United States and Britain since 1950*. Oxford: Oxford University Press.
- Olson, E. (2017). Personal identity. *The Stanford Encyclopedia of Philosophy*. E.N. Zalta (Ed.). Retrieved August 10, 2018 from <https://plato.stanford.edu/archives/sum2017/entries/identity-personal/>
- Ostapiuk, A. (2017a). Matematyzacja ekonomii – grzech pierworodny? Wieloaspektowa analiza wpływu i przyczyn. *Ekonomia XXI Wieku*, 13(1), 91–105.
- Ostapiuk, A. (2017b). Moralna ekonomia – *homo oeconomicus* jako istota pomagająca innym. *Ekonomia XXI Wieku*, 1(13), 70–90.
- Ostapiuk, A. (2018). W stronę dobrostanu. Działalność instytucji finansowych a nieracjonalność człowieka. *Ekonomia XXI Wieku*, 1(17), 82–101.
- Ostapiuk, A. (2019a). Droga ekonomii wolnej od wartościowania do epistemologicznej pychy. Użycie i nadużycie matematyki przez ekonomistów. *Zagadnienia Filozoficzne w Nauce*, (67), 153–202.
- Ostapiuk, A. (2019b). Human now versus human over time. When instrumental rationality and utility are not enough. *Panoeconomicus*, 66(5), 633–665.

- Ostapiuk, A. (2019c). Ekonomia neoklasyczna versus ekonomia behawioralna. Libertariański paternalizm a reforma emerytalna. *Annales. Ethics in Economic Life*, 22(1), 7–30.
- Oswald, A. J. (1997). Happiness and economic performance. *The Economic Journal*, 107(445), 1815–1831.
- Parfit, D. (1971). Personal identity. *The Philosophical Review*, 80(1), 3–27.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.
- Pauer-Studer, H. (2006). Identity, commitment and morality. *Journal of Economic Methodology*, 13(3), 349–369.
- Pavot, W., & Diener, E. (1993). The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research*, 28(1), 1–20.
- Perry, J. (1972). Can the self divide? *The Journal of Philosophy*, 69(16), 463–488.
- Perry, J. (1975). *Personal identity*. Berkeley: University of California Press.
- Pettit, P. (1996). Freedom as antipower. *Ethics*, 106(3), 576–604.
- Pettit, P. (2001). Symposium on Amartya Sen's philosophy: 1 Capability and freedom: A defence of Sen. *Economics & Philosophy*, 17(1), 1–20.
- Planck, M. (1950). *Scientific autobiography and other papers*. New York: Philosophical Library.
- Plato. (1961). Protagoras. In E. Hamilton, & H. Cairns (Eds.), *The collected dialogues of Plato* (pp. 308–352). Princeton: Princeton University Press.
- Plato. (1999). *Phaedrus*. The Project Gutenberg (B. Jowett, Trans.). Retrieved March 3, 2019 from <https://www.gutenberg.org/ebooks/1636>
- Plato. (2013). *Phaedo*. The Project Gutenberg (B. Jowett, Trans.). Retrieved March 3, 2019 from <https://www.gutenberg.org/ebooks/1658>
- Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
- Popper, K. (1983). *Realism and the aim of science*. Totowa, NJ: Rowman and Littlefield.
- Poundstone, W. (2008). *Gaming the Vote. Why elections aren't fair (and what we can do about it)*. New York: Hill and Wang.
- Pugmire, D. (1982). Motivated Irrationality. *Proceedings of the Aristotelian Society*, 56, 179–96.
- Pugno, M. (2004a). *The happiness paradox: A formal explanation from psycho-economics* (Working Paper no. 0501, Department of Economics, University of Trento).
- Pugno, M. (2004b). *Rationality and affective motivations: New ideas from neurobiology and psychiatry for economic theory?* (Discussion Paper no. 1, 2004, Department of Economics, University of Trento).
- Pugno, M. (2007). The subjective well-being paradox: A suggested solution based on relational goods. In L. Bruni, & P. Porta (Eds.), *Handbook on the economics of happiness* (pp. 263–289). Cheltenham: Edward Elgar.
- Putnam, H. (2002). *The collapse of the fact/value dichotomy and other essays*. Cambridge, MA: Harvard University Press.
- Putnam, H., & Walsh, V. (Eds.). (2011). *The end of value-free economics*. London: Routledge.
- Qizilbash, M. (2009). *Well-being, preference formation and the danger of paternalism* (Papers on Economics and Evolution, no. 0918, Max Planck Institute, University of Jena). Retrieved May 5, 2019 from <https://www.econstor.eu/bitstream/10419/32663/1/617152780.pdf>

- Qizilbash, M. (2011). Sugden's critique of Sen's capability approach and the dangers of libertarian paternalism. *International Review of Economics*, 58(1), 21–42.
- Qizilbash, M. (2012). Informed desire and the ambitions of libertarian paternalism. *Social Choice and Welfare*, 38(4), 647–658.
- Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60(1), 20–43.
- Quine, W. V. (1969). *Ontological relativity and other essays*. New York: Columbia University Press.
- Rabin, M. (1995). *Moral preferences, moral constraints, and self-serving biases* (Working Paper no. 95-241, University of California at Berkeley).
- Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, 36(1), 11–46.
- Radnitzky, G., & Bernholz, P. (Eds.). (1987). *Economic imperialism: The economic approach applied outside the field of economics*. New York: Paragon House Publishers.
- Rasmussen, D. C. (2006). Does 'bettering our condition' really make us better off? Adam Smith on progress and happiness. *American Political Science Review*, 100(3), 309–318.
- Ratajczak, M. (2005). Instytucjonalizm, wzbogacenie czy alternatywa ekonomii głównego nurtu? *Ekonomia i Prawo*, 1, 57–72.
- Ratajczak, M. (2011). Popularność ekonomii instytucjonalnej. Moda czy trwała zmiana? *Zeszyty Naukowe/Polskie Towarzystwo Ekonomiczne*, (9), 29–42.
- Ratajczak, M. (2014). Ekonomia i edukacja ekonomiczna w dobie finansyzacji gospodarki. *Ekonomista*, (2), 207–219.
- Ratajczak, M. (2017). Ekonomia dziś – ekonomia jutro. *Studia Oeconomica Posnaniensia*, 5(11), 20–34.
- Read, D. (2006). Which side are you on? The ethics of self-command. *Journal of Economic Psychology*, 27(5), 681–693.
- Read, D. (2007). Experienced utility: UTILITY theory from Jeremy Bentham to Daniel Kahneman. *Thinking & Reasoning*, 13(1), 45–61.
- Rebonato, R. (2014). A critical assessment of libertarian paternalism. *Journal of Consumer Policy*, 37(3), 357–396.
- Redman, D. (1993). *Economics and the philosophy of science*. Oxford: Oxford University Press.
- Reid, T. (1850). *Essays on the intellectual powers of man*. Retrieved May 10, 2018 from <https://play.google.com/store/books/details?id=54pHAAAIAAJ>
- Reiss, J. (2013). *Philosophy of economics: A contemporary introduction*. Abingdon: Routledge.
- Reiss, J. (2017). Fact-value entanglement in positive economics. *Journal of Economic Methodology*, 24(2), 134–149.
- Richardson, A. (2003). The geometry of knowledge: Lewis, Becker, Carnap and the formalization of philosophy in the 1920s. *Studies in History and Philosophy of Science Part A*, 34(1), 165–182.
- Rizzo, M. J., (2016). *Behavioral economics and deficient willpower: Searching for akrasia*. Retrieved from <https://ssrn.com/abstract=2731818>
- Robbins, L. (1932). *An essay on the nature & significance of economic science*. London: Macmillan.
- Robeyns, I. (2011). Book review of Martha Nussbaum's 'Creating capabilities'. *Notre Dame Philosophy Reviews*. Retrieved May 29, 2018 from <https://ndpr.nd.edu/news/creating-capabilities-the-human-development-approach-2/>

- Robeyns, I. (2016). The capability approach. *The Stanford Encyclopedia of Philosophy*, E. Zalta (Ed.). Retrieved October 12, 2018 from <https://plato.stanford.edu/archives/win2016/entries/capability-approach/>
- Rodrik, D. (2015). *Economics rules: The rights and wrongs of the dismal science*. New York: W.W. Norton & Co.
- Róna, P., & Zsolnai, L. (Eds.). (2017). *Economics as a moral science*. Springer International Publishing.
- Rorty, A. (Ed.). (1976) *The identities of persons*. Berkeley: University of California Press.
- Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.
- Rosenberg, A. (1992). *Economics-mathematical politics or science of diminishing returns?* Chicago: University of Chicago Press.
- Ross, D. (2005). *Economic theory and cognitive science*. Cambridge, MA: The MIT Press.
- Rothbard, M. N. (1978). *For a new liberty: The libertarian manifesto*. Auburn, AL: Ludwig von Mises Institute.
- Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, 12(3), 185–190.
- Rousseau, J. (1979). *Emile or on education* (A. Bloom, Trans.). New York: Basic Books.
- Royal Swedish Academy of Sciences. (1993). The Nobel Memorial Prize in Economics 1992. Press Release from the Royal Swedish Academy of Sciences. *Scandinavian Journal of Economics*, 95(1), 1–5.
- Rubinstein, A. (2008). Comments on neuroeconomics. *Economics & Philosophy*, 24(3), 485–494.
- Runciman, W. (1968). Sociological evidence and political theory. In M. Brodbeck (Ed.), *Readings in the philosophy of the social sciences* (pp. 561–571). New York: Macmillan.
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52(1), 141–166.
- Sady, W. (2013). *Spór o racjonalność naukową: od Poincarégo do Laudana*. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Sagoff, M. (1986). Values and preferences. *Ethics*, 96(2), 301–316.
- Salmon, W. C. (1989). *Four Decades of Scientific Explanation*. Pittsburgh, PA: University of Pittsburgh Press.
- Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, 4(2), 155–161.
- Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61–71.
- Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. *Economica*, 15(60), 243–253.
- Samuelson, P. A. (1950). The problem of integrability in utility theory. *Economica*, 17(68), 355–85.
- Samuelson, P. A. (1952). Economic theory and mathematics – an appraisal. *American Economic Review*, 42(2), 56–66.
- Samuelson, P. A. (1992). My life philosophy: Policy credos and working ways. In M. Szenberg (Ed.), *Eminent economists: Their life philosophies* (pp. 236–247). Cambridge: Cambridge University Press.

- Samuelson, P. A., & Nordhaus, W. D. (2001). *Economics*. New York: McGraw-Hill.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59.
- Sandel, M. J. (1998). *Liberalism and the limits of Justice*. Cambridge: Cambridge University Press.
- Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality*, 61(3), 317–342.
- Sartre, J. P. (1992). *Being and nothingness: A phenomenological essay on ontology*. New York: Washington Square Press.
- Sartre, J. P. (2007). *Existentialism is a humanism*. New Haven, CT: Yale University Press.
- Sartre, J. P. (2015). *Sketch for a theory of the emotions*. London: Routledge.
- Schafer, R. (1976). *A new language for psychoanalysis*. New Haven, CT: Yale University Press.
- Schelling, T. C. (1978). Egonomics, or the art of self-management. *The American Economic Review*, 68(2), 290–294.
- Schelling, T. C. (1980). The intimate contest for self-command. *The Public Interest*, 60, 94–116.
- Schelling, T. C. (1984). Self-command in practice, in policy, and in a theory of rational choice. *The American Economic Review*, 74(2), 1–11.
- Schelling, T. C. (1985). Enforcing rules on oneself. *Journal of Law, Economics and Organization*, 1(2), 357–374.
- Schelling, T. C. (1996). Coping rationally with lapses from rationality. *Eastern Economic Journal*, 22(3), 251–269.
- Scheuer, B. (2015). *Metodologia ekonomii w perspektywie konstruktywistycznej*. Wrocław: Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.
- Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, 9(5), 340–346.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I: Detection, search, and attention. *Psychological Review*, 84(1), 1–66.
- Schnellenbach, J. (2016). A constitutional economics perspective on soft paternalism. *Kyklos*, 69(1), 135–156.
- Schooler, J., Ariely, D., & Loewenstein, G. (2003). The pursuit and assessment of happiness can be self-defeating. In I. Brocas, & J. Carrillo (Eds.), *The psychology of economic decisions*, vol. 1: *Rationality and well-being* (pp. 41–70). New York: Oxford University Press.
- Schwartz, B. (1990). The creation and destruction of value. *American Psychologist*, 45(1), 7–15.
- Schwartz, B. (2000). Self-determination: The tyranny of freedom. *American Psychologist*, 55(1), 79–88.
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Harper Collins.
- Schweitzer, A. (1970). Typological method in economics: Max Weber's contribution. *History of Political Economy*, 2(1), 66–96.
- Scitovsky, T. (1976). *The joyless economy: The psychology of human satisfaction*. Oxford: Oxford University Press.
- Scollon, C. N., Prieto, C. K., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, 4(1), 5–34.

- Screpanti, E., & Zamagni, S. (1993). *An outline of the history of economic thought*. Oxford: Clarendon Press.
- Sedláček, T. (2011). *Economics of good and evil: The quest for economic meaning from Gilgamesh to Wall Street*. Oxford, New York: Oxford University Press.
- Seligman, M. (1994). *What you can change and what you can't*. New York: Knopf.
- Sen, A. K. (1970). *Collective choices and social welfare*. Edinburgh: Oliver & Boyd.
- Sen, A. K. (1971). Choice functions and revealed preference. *The Review of Economic Studies*, 38(3), 307–317.
- Sen, A. K. (1973). Behaviour and the concept of preference. *Economica*, 40(159), 241–259.
- Sen, A. K. (1974). Choice, orderings, and morality. In S. Korner (Ed.), *Practical reason* (pp. 54–67). Oxford: Blackwell.
- Sen, A. K. (1977). Rational fools: A critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs*, 6(4), 317–344.
- Sen, A. K. (1979). Utilitarianism and welfarism. *The Journal of Philosophy*, 76(9), 463–489.
- Sen, A. K. (1982). Liberty as control: An appraisal. *Midwest Studies in Philosophy*, 7(1), 207–221.
- Sen, A. K. (1983). Liberty and social choice. *Journal of Philosophy*, 80(1), 5–28.
- Sen, A. K. (1985a). *Commodities and capabilities*. Amsterdam: North-Holland.
- Sen, A. K. (1985b). The standard of living: Concepts and critiques. In G. Hawthorn (Ed.), *The Tanner lectures* (pp. 1–51). Cambridge: Clare Hall.
- Sen, A. K. (1985c). Well-being, agency and freedom: The Dewey lectures. *The Journal of Philosophy*, 82(4), 169–223.
- Sen, A. K. (1987). *On ethics & economics*. Malden, MA: Blackwell.
- Sen, A. K. (1988). Freedom of choice: Concept and content. *European Economic Review*, 32(2-3), 269–294.
- Sen, A. K. (1990). Gender and cooperative conflicts. In I. Tinker (Ed.), *Persistent inequalities* (pp. 123–149). New York: Oxford University Press.
- Sen, A. K. (1992). *Inequality reexamined*. Cambridge, MA: Harvard University Press.
- Sen, A. K. (1993a). Internal consistency of choice. *Econometrica*, 61(3), 495–521.
- Sen, A. K. (1993b). Capability and well-being. In M. Nussbaum, & A. Sen (Eds.), *The quality of life* (pp. 30–53). Oxford: Oxford University Press.
- Sen, A. K. (1997). Maximization and the act of choice. *Econometrica*, 65(4), 745–779.
- Sen, A. K. (1999). *Development as freedom*. New York: Knopf.
- Sen, A. K. (2002). *Rationality and freedom*. Cambridge, MA: Belknap Press.
- Sen, A. K. (2006). Reason, freedom and well-being. *Utilitas*, 18(1), 80–96.
- Sen, A. K. (2009). *The Idea of Justice*. London: Penguin Books.
- Sen, A. K. (2010). Adam Smith and the contemporary world. *Erasmus Journal for Philosophy and Economics*, 3(1), 50–67.

- Shoemaker, D. (2016). Personal identity and ethics. *The Stanford Encyclopedia of Philosophy*. E. Zalta (Ed.). Retrieved July 8, 2018 from <https://plato.stanford.edu/archives/win2016/entries/identity-ethics/>
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7(4), 269–285.
- Shoemaker, S. (1984). Personal identity: A materialist's account'. In S. Shoemaker, & R. Swinburne (Eds.), *Personal identity* (pp. 67–132). Oxford: Blackwell.
- Shoemaker, S. (1997). Self and substance. In J. Tomberlin (Ed.), *Philosophical perspectives 11: Mind, causation and world* (pp. 288–289). Boston, Oxford: Blackwell.
- Shoemaker, S. (1999). Self, body, and coincidence. *Proceedings of the Aristotelian Society*, 73(1), 287–306.
- Siderits, M. (2019). Buddha. *The Stanford Encyclopedia of Philosophy*. Retrieved February 3, 2019 from <https://plato.stanford.edu/archives/spr2019/entries/buddha/>
- Simon, H. A. (1957). *Administrative behavior: A study of decision-making processes in administrative organization* (2nd ed). New York: Macmillan.
- Skousen, M. (1997). The perseverance of Paul Samuelson's economics. *Journal of Economic Perspectives*, 11(2), 137–152.
- Smith, A. (1759). *The theory of moral sentiments*. Library of Economics and Liberty. Retrieved February 2, 2018 from <http://www.econlib.org/library/Smith/smMS.html>
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. The Project Gutenberg. Retrieved April 2, 2019 from <http://www.gutenberg.org/ebooks/3300>
- Smith, R. H., Diener, E., & Wedell, D. H. (1989). Intrapersonal and social comparison determinants of happiness: A range-frequency analysis. *Journal of Personality and Social Psychology*, 56(3), 317–325.
- Solomon, R. (1980). Emotions and choice. In A. Rorty (Ed.), *Explaining the emotions* (pp. 251–282). Berkeley, Los Angeles: University of California Press.
- Sondaz: Polacy zadowoleni z zakazu palenia. (2010). Retrieved May 10, 2016 from <http://www.nton.pl/polska-i-swiat/art/4170131,sondaz-polacy-zadowoleni-z-zakazu-palenia,id,t.html>
- Sophie's choice. (n.d.). Oxford dictionary. Retrieved May 3, 2019 from https://en.oxforddictionaries.com/definition/sophie's_choice
- Spiegel, H. (1987). Jacob Viner (1892–1970). In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics*, vol. 4 (pp. 812–814). London: Palgrave Macmillan.
- Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Brookings Papers on Economic Activity*, 39(1), 1–102.
- Stigler, G. J. (1950). The development of utility theory. I. *Journal of Political Economy*, 58(4), 307–327.
- Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2), 76–90.
- Stiglitz, J. E., Sen, A., & Fitoussi, J. P. (2010). *Mismeasuring our lives: Why GDP doesn't add up*. New York: The New Press.
- Stillman, T. F., Baumeister, R. F., & Mele, A. R. (2011). Free will in everyday life: Autobiographical accounts of free and unfree actions. *Philosophical Psychology*, 24(3), 381–394.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8(3), 220–247.

- Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology*, 18(5), 429–442.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165–180.
- Stroud, S., & Tappolet, C. (Eds.). (2003). *Weakness of will and practical irrationality*. Oxford: Clarendon Press.
- Stutzer, A. (2004). The role of income aspirations in individual happiness. *Journal of Economic Behavior & Organization*, 54(1), 89–109.
- Styron, W. (1992). *Sophie's choice*. Vintage [Reissue edition].
- Sugden, R. (1991). Rational choice: A survey of contributions from economics and philosophy. *The Economic Journal*, 101(407), 751–785.
- Sugden, R. (2000). Credible worlds: The status of theoretical models in economics. *Journal of Economic Methodology*, 7(1), 1–31.
- Sugden, R. (2004). The opportunity criterion: Consumer sovereignty without the assumption of coherent preferences. *American Economic Review*, 94(4), 1014–1033.
- Sugden, R. (2006). What we desire, what we have reason to desire, whatever we might desire: Mill and Sen on the value of opportunity. *Utilitas*, 18(1), 33–51.
- Sugden, R. (2008a). Capability, happiness and opportunity. In L. Bruni, F. Comim, & M. Pugno (Eds.), *Capabilities and happiness* (pp. 299–322). Oxford: Oxford University Press
- Sugden, R. (2008b). Why incoherent preferences do not justify paternalism. *Constitutional Political Economy*, 19(3), 226–248.
- Sugden, R. (2009a). On nudging: A review of “Nudge: Improving decisions about health, wealth and happiness” by Richard H. Thaler and Cass R. Sunstein. *International Journal of the Economics of Business*, 16(3), 365–373.
- Sugden, R. (2009b). Credible worlds, capacities and mechanisms. *Erkenntnis*, 70(1), 3–27.
- Sugden, R. (2010). Opportunity as mutual advantage. *Economics & Philosophy*, 26(1), 47–68.
- Sugden, R. (2015). Looking for a psychology for the inner rational agent. *Social Theory and Practice*, 41(4), 579–598.
- Sugden, R. (2017). Do people really want to be nudged towards healthy lifestyles? *International Review of Economics*, 64(2), 113–123.
- Sugden, R. (2018). *The community of advantage: A behavioural economist's defence of the market*. New York, Oxford: Oxford University Press.
- Sunstein, C. R. (2013a). *Simpler: The future of government*. New York: Simon & Schuster.
- Sunstein, C. R. (2013b). The Storrs lecture: Behavioral economics and paternalism. *The Yale Law Journal*, 122(7), 1826–1899.
- Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 70(4), 1159–1202.
- Sunstein, C. R., & Thaler, R. H. (2005). Libertarian paternalism. In C. R. Sunstein (Ed.), *Laws of fear: Beyond the precautionary principle* (pp. 175–203). Cambridge: Cambridge University Press.
- Swedberg, R. (1990). *Economics and sociology*. Princeton, NJ: Princeton University Press.

- Tambolo, L. (2015). A tale of three theories: Feyerabend and Popper on progress and the aim of science. *Studies in History and Philosophy of Science Part A*, 51, 33–41.
- Taylor, C. (1992). *The ethics of authenticity*. Cambridge, MA: Harvard University Press.
- Tenenbaum, S. (1999). The judgement of a weak will. *Philosophy and Phenomenological Research*, 59(4), 875–911.
- Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1(1), 39–60.
- Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199–214.
- Thaler, R. H. (2000). From homo economicus to homo sapiens. *Journal of Economic Perspectives*, 14(1), 133–141.
- Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York: W.W. Norton & Company.
- Thaler, R. H. (2016). Behavioral economics: Past, present, and future. *American Economic Review*, 106(7), 1577–1600.
- Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy*, 89(2), 392–406.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, 93(2), 175–179.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
- The Declaration of Independence*. (1776). Retrieved March 3, 2019 from <http://www.ushistory.org/declaration/document/>
- Tichý, P. (1974). On Popper's definitions of verisimilitude. *The British Journal for the Philosophy of Science*, 25(2), 155–160.
- Truc, A. (2018). Is 'new' behavioral economics 'mainstream'? *Journal of Economic Methodology*, 25(1), 83–104.
- Turner, B. (2010). Homo economicus: Alive and diversified in the economic research program. *Syracuse University Honors Program Capstone Projects*. 399. Retrieved May 1, 2017 from https://surface.syr.edu/cgi/viewcontent.cgi?article=1396&context=honors_capstone
- Tversky, A., & Griffin, D. (1991). Endowment and contrast in judgments of well-being. In R. Zeckhauser (Ed.), *Strategy and choice* (pp. 297–318). Cambridge, MA: The MIT Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Unger, P. (1990). *Identity, consciousness and value*. New York: Oxford University Press.
- Unger, P. (2000). The survival of the sentient. *Philosophical Perspectives*, 14, 325–348.
- Urbina, D. A., & Ruiz-Villaverde, A. (2019). A critical review of homo economicus from five approaches. *American Journal of Economics and Sociology*, 78(1), 63–93.
- Urry, H. L., Nitschke, J. B., Dolski, I., Jackson, D. C., Dalton, K. M., Mueller, C. J., ... & Davidson, R. J. (2004). Making a life worth living: Neural correlates of well-being. *Psychological Science*, 15(6), 367–372.

- Van Dun, F. (1986). Economics and the limits of value-free science. *Reason Papers*, (11), 17–32.
- Van Staveren, I. (2001). *The values of economics: An Aristotelian perspective*. London: Routledge.
- Van Staveren, I. (2007). Beyond utilitarianism and deontology: Ethics in economics. *Review of Political Economy*, 19(1), 21–35.
- Vickery, W. S. (1964). *Microstatics*. New York: Harcourt, Brace and World.
- Vromen, J. (2007). In praise of moderate plurality. In J. Groenewegen (Ed.), *Teaching pluralism in economics* (pp. 64–94). Cheltenham: Edward Elgar.
- Wallace, R. J. (1999). Addiction as defect of the will: Some philosophical reflections. *Law and Philosophy*, 18(6), 621–654.
- Watson, G. (1977). Skepticism about weakness of will. *The Philosophical Review*, 86(3), 316–339.
- Weber, M. (1949). *The methodology of the social sciences* (E. Shils, & H. Finch, Trans.). Glencoe: The Free Press.
- Weber, M. (2012). Science as a profession and a vocation. In H. Bruun, & S. Whimster (Eds.), *Max Weber: Collected methodological writings*. London, New York: Routledge.
- Wedgwood, R. (2003). Choosing rationally and choosing correctly. In S. Stroud, & C. Tappolet (Eds.), *Weakness of will and practical irrationality* (pp. 201–229). Oxford: Clarendon Press.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: The MIT Press.
- Weijers, D. (2011). Hedonism. *The Internet Encyclopedia of Philosophy*. Retrieved May 13, 2019 from <https://www.iep.utm.edu/hedonism/>
- Weintraub, E. R. (1985). *General equilibrium analysis: Studies in appraisal*. Cambridge: Cambridge University Press.
- Weintraub, E. R. (2002). *How economics became a mathematical science*. Durham, London: Duke University Press.
- Wells, T. (2013). Adam Smith on morality and self-interest. In C. Luetge (Ed.), *Handbook of the philosophical foundations of business ethics* (pp. 281–296). Dordrecht: Springer.
- White, M. D. (2004). Can homo economicus follow Kant’s categorical imperative? *The Journal of Socio-Economics*, 33(1), 89–106.
- White, M. D. (2006). Multiple utilities and weakness of will: A Kantian perspective. *Review of Social Economy*, 64(1), 1–20.
- White, M. (2011). *Kantian ethics and economics: Autonomy, dignity, and character*. Stanford: Stanford University Press.
- White, M. D. (2013). *The manipulation of choice: Ethics and libertarian paternalism*. New York: Palgrave Macmillan.
- Wilkin, J. (2016). *Instytucjonalne i kulturowe podstawy gospodarowania: humanistyczna perspektywa ekonomii*. Warszawa: Wydawnictwo Naukowe Scholar.
- Williams, B. (1970). Deciding to believe. In B. Williams (Ed.), *Problems of the self* (pp. 136–51). Cambridge: Cambridge University Press.
- Williams, B. (1973). A critique of utilitarianism. In J. Smart, B. Williams (Eds.), *Utilitarianism: For and against* (pp. 77–155). Cambridge: Cambridge University Press.

- Wilson, C. M., & Price, C. W. (2010). Do consumers switch to the best supplier? *Oxford Economic Papers*, 62(4), 647–668.
- Wilson, T. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Belknap Press/Harvard University Press.
- Witztum, A. (2007). Ethics and the science of economics: Robbins's enduring fallacy. In F. Cowell, & A. Witztum (Eds.), *Lionel Robbins' essay on the nature and significance of economic science. 75th anniversary conference proceedings* (pp. 57–85). London: School of Economics and Political Science.
- Wolin, S. S. (1981). Max Weber: Legitimation, method, and the politics of theory. *Political Theory*, 9(3), 401–424.
- Wong, S. (2006). *Foundations of Paul Samuelson's revealed preference theory: A study by the method of rational reconstruction*. London: Routledge.
- Yuengert, A. M. (2000). *The positive-normative distinction before the fact-value distinction*. Malibu, CA: Pepperdine University.
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–75.
- Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. Forgas (Ed.), *Feeling and thinking* (pp. 31–58). Cambridge: Cambridge University Press.
- Zak, P. J. (2004). Neuroeconomics. *Transactions of the Royal Philosophical Society B*, 359(1451), 1737–1748.
- Zboroń, H. (2009a). Konstruktywizm społeczny – nowe teoretyczne podejście w badaniach nad gospodarką. *Prakseologia*, (149), 63–87.
- Zboroń, H. (2009b). *Teorie ekonomiczne w perspektywie poznawczej konstruktywizmu społecznego*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Zboroń, H. (2013). Dyskurs metodologiczny we współczesnej ekonomii – próba dekonstrukcji podejścia modernistycznego. *Studia Ekonomiczne*, 1(76), 13–45.