UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

ÉCOLE DOCTORALE SCIENCES HUMAINES ET SOCIALES (555)

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

DOCTOR OF PHILOSOPHY (PhD) IN MANAGEMENT FROM NEOMA BUSINESS SCHOOL

Discipline* : *SCIENCES ECONOMIQUES

Présentée et soutenue par

Nestor LOVERA NIETO

Le 17 novembre 2022

Elucidating the Role of Value Judgments in Normative Economics

Thèse dirigée par M. Cyril HÉDOIN

Co-dirigée par M. Stéphane LHUILLERY

JURY

Mme Herrade IGERSHEIM,	Directrice de Recherche,	Université de Strasbourg,	**Présidente du jury**
M. Cyril HÉDOIN,	Professeur des Universités,	Université de Reims Champagne-Ardenne,	**Directeur de thèse**
M. Stéphane LHUILLERY,	Full Professor,	Neoma Business School,	**Co-Directeur**
Mme Constanze BINDER,	Associate Professor,	Erasmus University Rotterdam,	**Rapporteur**
M. Marc FLEURBAEY,	Directeur de Recherche,	Paris School of Economics,	**Rapporteur**
M. Rouslan KOUMAKHOV,	Full Professor,	Neoma Business School,	**Examinateur**
M. Fouad BEN-ABDELAZIZ,	Full Professor,	Neoma Business School,	**Examinateur**
M. Richard ARENA,	Professeur émérite,	Université Côte d'Azur,	**Examinateur**

Acknowledgements

I would like to thank a large number of people who contributed directly or indirectly to the writing of this dissertation. From the beginning I knew that it would be one of the most important challenges of my life, and today I can confirm that it is the most demanding academic work I have ever done. Without a doubt, I was not alone to face this challenge. Many people gave me their support, academically, professionally and personally. I apologize that I have not mentioned all of them in these acknowledgments.

I would like to start by thanking my supervisor Cyril Hédoin. Our frequent discussions were crucial in the development of the dissertation. These discussions over four years allowed me to train as a researcher. His comments and questions about my research led me to reflect on my work and to explore new possibilities. His open-mindedness allowed me to have the liberty to choose and explore the topics of my predilection. His benevolence was important in giving me confidence and making me feel that I had the capacity to face this challenge.

I also thank my Neoma Business School supervisor Stéphane Lhuillery. His interesting remarks about the relevance of my dissertation topic were essential to strengthen my research question. His vision as a non-specialist in the field of philosophy of economics was especially useful to make me reflect on the scope of my dissertation.

My gratitude is addressed to the committee members of the thesis for accepting to evaluate my research. I want to thank Herrade Igersheim who was one of the two members of my PhD follow-through committee (with Jean-Sébastien Gharbi). Her comments and suggestions were always of great help to improve my work. Furthermore, she was my tutor for commenting my two working papers at the 23rd and 24th ESHET Summer Schools in History of Economic Thought, Economic Philosophy and Economic History in Paris and in Strasbourg. Our discussions in the summer school tutorials were thought-provoking and helped me to advance in my research.

I met Constanze Binder at the Workshop “Social Justice in a Complex World” organized by the economic philosophy section of the REGARDS research center on November 22, 2019. I discovered her book “Agency, Freedom and Choice” (2019), which was really in line with my dissertation. So, I wrote a review article of the book. It also allowed me to participate in the symposium on her book at the study group of The Dutch Research School of Philosophy (OZSW) and the Erasmus Institute for Philosophy on Economics (EIPE). I am truly thankful for her remarks on my comments on her book. The book was a source of inspiration for the second chapter of this dissertation.

I also want to thank Fouad Ben-Abdelaziz who, together with Philippe Odou, encouraged me to enroll at Neoma Business School’s PhD program. My gratitude goes to Richard Arena, Marc Fleurbaey and Rouslan Koumakhov for taking an interest in my work and accepting to be part of the committee to evaluate my dissertation.

I express my gratitude to Marion Gaspard, Rebeca Gomez Betancourt and Michaël Assous, my former professors at the University of Lyon 2 where I obtained my master's degree in the History of Economic Thought, for encouraging me to continue my PhD studies.

I thank the REGARDS research center for allowing me to do my dissertation under the best possible conditions. A special thanks goes to Jean-Sébastien Gharbi with whom I had thought-provoking discussions about my dissertation and for his helpful advice on my short teaching career. I am truly thankful to Pierre Van Zyl for his careful reading. His corrections were essential to improve the understanding of the final manuscript. I want to thank my young colleagues Agnès Penasse, Aurore Sivignon, Manu Navarro, Baptiste Privé, Dorian Marchais and Quentin Oget with whom I shared my dissertation years. Without them the road would have been less pleasant.

I am very grateful to my family. Without their unconditional support this dissertation would never have been written. A special thanks goes to my parents, Edelmira Nieto and César Lovera, who supported my project of studying in France without hesitation. Many thanks also go to my aunt, Nieves Nieto, who has always been in the background to tell me that everything is possible and that you just have to persevere to get good results. I think of my grandmother who passed away during the writing of the dissertation. Her tenacity and fighting spirit were my inspiration in the moments when I doubted my ability to face this challenge.

Finally, I have to thank my friends, or as my wise mother says, "the family that life gave you." I thank Loipa Muniz who gave me her unconditional support to prepare the audition to obtain

the doctoral contract, without whose help I would not have obtained it. Many thanks go to Janine Guimiot who was always present to listen to me and give me encouragement when I needed it. I am grateful to Patrice Farrugia who made my daily life so much easier during the writing of this dissertation. Thanks to Claude Durif, my dear Bala, who welcomed me as a member of her family during my master studies in Lyon. I also thank other friends who indirectly contributed to the success of this dissertation: Carely, Nelly, and Elsy for always believing in me and giving me encouragement when I needed it; Christian for reminding me that it is important to take a break and pursue a few other passions.

Contents

List of Tables

General Introduction

Normative Economics as a Science

The distinction between positive economics and normative economics was first made by John Neville Keynes in *The Scope and Method of Political Economy* (1917). Positive economics is defined as "a body of systematized knowledge concerning what is," while normative economics is defined 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" (Keynes 1890 [1917], 34–5). However, it is important to note that Keynes makes a threefold distinction of economic inquiries. In addition to positive economics and normative economics, he identifies the art of political economy defined as "a system of rules for the attainment of a given end […] which seeks to formulate economic precepts" (Keynes 1890 [1917], 35–6). The importance of this distinction lies in giving positive economics and normative economics the same status as a science.[1] This is the consequence of separating the normative component from the artistic component, marking a difference with authors such as Mill (1844 [1967]) and Senior (1852 [1966]) who included the normative component in what they identified as art.[2] The Keynesian distinction allows us to maintain a separation between positive economics and normative economics without minimizing the scientific character of normative economics.[3]

However, Keynes's position regarding the status of normative economics as science is not shared by the methodologist Lionel Robbins, whose analysis of the positive-normative

[1] "A department of knowledge does not necessarily belong to the category of art, as distinguished from science, simply because it is concerned with what ought to be. Logic and ethics are both of them sciences although they are concerned with right reasoning and right conduct respectively" (Keynes 1890 [1917], 35).
[2] See Mongin (2018) who emphasizes that the originality of Keynes's distinction lies in the fact that he separates the normative component of economics from the artistic component of economics.
[3] See Badiei (2022) who defends the distinction between positive economics and normative economics taking Keynes's distinction as a reference. Badiei analyzes the hypothesis that the distinction between normative economics and positive economics is necessary. He shows that it is problematic to reduce the importance of normative economics with respect with positive economics.

distinction has been extensively studied. Robbins appeals to Hume's thesis that one cannot deduce an "ought" from an "is" to establish a strict separation between positive economics and normative economics. "Propositions involving the verb 'ought' are different in kind from propositions involving the verb 'is.' And it is difficult to see what possible good can be served by not keeping them separate, or failing to recognize their essential difference" (Robbins 1935 [1984], 149).[4] While Hume's influence on Robbins's work is well known, the influence of Max Weber is not. Mongin (2018) points out that one of the elements that Robbins takes from Weber are the conceptions of judgment of fact and judgment of value, which enrich the positive-normative distinction. Robbins states that in order to understand the concepts of preference and choice in economics, the objectivity proposed by Weber is important.

> All that the "objective" (that is to say, the *wertfrei*, to use Max Weber's phrase) explanation of conduct involves is the consideration of certain data, individual valuations, etc., which are not merely physical in character. The fact that such data are themselves of the nature of judgments of value does not necessitate that they should be valued as such (Robbins 1935 [1984], 90).

This implies that although economists can consider value judgments that may be associated with moral, political or ideological convictions of the individuals under study, it is possible for economists to make judgments of fact about them. In this way, Robbins proposes a separation by the juxtaposition of normative economics and positive economics. But such a juxtaposition is not made by considering positive economics and normative economics under the same status as science. Robbins grants the status of science to positive economics, while normative economics is about ethics and is therefore far from being a pure application of science. He assumes that economics is exclusively positive. This position has influenced the debate on the distinction between positive and normative economics for a few decades.

Milton Friedman supports the distinction proposed by Robbins "Positive economics is in principle independent of any particular ethical position or normative judgment. In short, positive economics is, or can be, an "objective" science, in precisely the same sense as any of the physical sciences" (Friedman 1953, 4). He argues that differences about economic policy are the result of a disagreement about the consequences that positive economics associates with diverse possible actions rather than a disagreement about the basic values that normative

[4] This passage has been the subject of much debate. On the one hand, Putnam (2002) and Davis (2013) claim that Robbins assumes the logical positivists' view to analyze the positive-normative distinction in economics. On the other hand, Colander and Su (2015) argue that Robbins is influenced by what they call the Mill-Keynes tradition, which considers that ethics plays an important role in economic policy analysis.

economics associates with those consequences or actions.[5] Friedman adds that in the case of normative disagreements "men can ultimately only fight" (Friedman 1953, 5). This leads him to say that: "a consensus on 'correct' economic policy depends much less on the progress of normative economics proper than on the progress of a positive economics yielding conclusions that are, and deserve to be, widely accepted" (Friedman 1953, 6). This implies that for him normative economics is subordinated to positive economics.[6]

After having identified two ways by which positive and normative economics can be separated, it is important to consider Putnam's (2002) view that positive and normative economics are entangled. Putnam (2002) and Putnam and Walsh (2012) reject the logical positivist position expressed by the fact/value dichotomy. This position considers that empirical science and logic/mathematics are the only types of meaningful discourse, while ethical, metaphysical and aesthetic judgments are not meaningful (Hands 2012). Putnam proposes to give up the dichotomy between facts and values and argues that they are entangled. This can be explained in the social sciences and more specifically in economics by the existence of thick concepts, whose content is simultaneously descriptive and normative.[7] For this reason, Putnam affirms that: "Certainly the world we inhabit when we describe the world for the purposes the economist is interested in is not describable in 'value-neutral' terms. Not without throwing away the most significant facts along with the 'value judgments'" (Putnam and Walsh 2012, 112). Putnam wants to show that it is possible in economic theory to "combine both the best mathematical models of an economy and the best ethical reflection" (Ibid, 112).[8]

Although we have not made an exhaustive analysis of the debate on the distinction between positive and normative economics, we can say that the strict separation between positive and normative economics has led some to consider that the normative character of economics is not part of science. Hence the importance of considering the distinction made by Keynes, who gives the same status of science to both positive economics and normative economics. However, his

[5] Mongin (2018) considers that Friedman's analysis of the economic policy controversies is important to understand Friedman's position on the scope of normative economics.

[6] See Chapter 4 of Badiei (2022) who makes a critical study of Friedman's position regarding the distinction between positive and normative economics.

[7] Putnam tries to explain what a thick concept is through the example of the word "cruel":
"The word "cruel" obviously – or at least it is obvious to most people, even if it is denied by some famous defenders of the fact/value dichotomy – has normative and, indeed, ethical uses. If one asks me what sort of person my child's teacher is, and I say, "He is very cruel," I have both criticized him as a teacher and criticized him as a man […] Yet "cruel" can also be used purely descriptively, as when a historian writes that a certain monarch was exceptionally cruel, or that the cruelties of the regime provoked a number of rebellions" (Putnam 2002, 34–5).

[8] In the same vein as Putnam, Hands (2012) considers that positive and normative economics are entangled and that the difference between them cannot constitute a strict dichotomy. Davis (2013) criticizes the value neutrality view held by most economists based on the positive-normative dichotomy.

definition of normative economics is limited to his definition of political economy. Keynes considers that political economy is mainly limited to studying the phenomenon of wealth, so that normative economics is restricted to evaluating diverse social situations with respect to the creation, appropriation, accumulation, and redistribution of wealth (Keynes 1890 [1917], 2).

As pointed out by Badiei (2022) and Badiei, Campagnolo, and Grivaux (2022), after the marginalist revolution, economic science came to be considered as a rational decision science rather than a wealth science. Therefore, the definition of normative economics proposed by Keynes does not respond to the evolution of ideas in economics. For this reason, we assume the following definition of normative economics: it is concerned with investigating the methods and criteria to evaluate the desirable or undesirable character of economic states of affairs and also with proposing the way in which these methods and criteria can be implemented in economics (Badiei 2022; Mongin 2002; 2006a).

Research Question and Main Hypothesis

We consider that the efforts to reappraise the role of normative economics made by authors such as Putnam, Hands, Davis, among others, should be maintained and strengthened. For this reason, the objective of this dissertation is to determine the role that value judgments play in the study of normative economics. In other words, we want to determine whether normative economics makes value judgments or is only related to value judgments made by economic agents. This leads us to ask ourselves the following questions: (1) To what extent are economists willing to make value judgments for themselves when studying normative economics? (2) How do economists take into account the value judgments of economic agents in the study of normative economics? (3) How do normative economists study and characterize a value (freedom) in relation with the value judgments of economic agents (as represented by their preferences)? (4) To what extent are value judgments important in the elaboration and interpretation of normative economic theories?

In order to achieve the objective of this dissertation we focus on studying the way economists practice normative economics. In this way, we first propose to establish a theoretical framework that allows us to distinguish the value judgments of economists and economic agents. Second, we study freedom of choice as a normative criterion, which leads us to examine how economists associate freedom with the preferences of individuals in order to define and measure individual

freedom. Third, we take Kenneth Arrow's impossibility theorem as a case study to show that Arrow cannot elaborate and interpret his impossibility result without making value judgments.

The main contribution of this dissertation will be to try and show that economists cannot dispense with value judgments in the study of normative economics. Value judgments are thus essential not only to choose the methods and criteria that will be useful for evaluating the desirable or undesirable character of economic situations, but also when formulating prescriptions for the implementation of these methods and criteria. The importance of this contribution lies in the fact that even today some economists continue to consider normative economics as part of positive economics so that value judgments are outside the scope of economics. For instance, Gul and Pesendorfer (2008) assert that "standard welfare economics functions as part of positive economics" (Ibid, 5), which leads them to state that "normative statements […] are used to define new positive questions […]" (Ibid, 25).[9]

In the following two sections we highlight the different forms that normative economics has taken throughout history. This will allow us to identify the different problems and theoretical answers of normative economics. We will give a definition of value judgments. At first, we refer to the philosophy of science to define value judgments. Then we examine how economists have defined value judgments. The idea is to see if there is a dialogue between the philosophy of science and economics regarding value judgments.

The Different Branches of Normative Economics

The first form of normative economics dates from the beginning of the 20th century. This corresponds to "economics of welfare", as Pigou (1920) called it, which is inspired by utilitarianism. This first form of welfare economics states the welfare conditions of the market in terms of Paretian optimality, and it bases its recommendations on interpersonal comparisons of utility. In other words, the idea is to calculate the well-being of each individual and arbitrate between the gains in well-being of some and the loss of well-being of others. Thus the prescriptions given by the models of this first form of normative economics are based on the cardinality of preferences and the possibility of making interpersonal comparisons of utility (Baujard 2017).

[9] Gul and Pesendorfer consider that standard welfare economics "tries to evaluate how economic institutions mediate […] behavior of agents" (Ibid, 5).

Interpersonal comparisons of utility were however strongly criticized, leading economists to refrain from using them. Robbins 1935 [1984] expressed one of the most important criticisms against interpersonal comparisons of utility. He considers that the "comparison of the different satisfactions of different individuals involves judgments of value rather than judgments of fact, and that such judgments are beyond the scope of positive science" (Robbins 1935 [1984], xxxiv). [10] Another important criticism leveled at interpersonal comparisons of utility refers to the operational meaning of interpersonal comparisons. Baujard (2017) states that "comparisons only have operational meaning if the differences, relationships, or levels of utility have meaning themselves, and if they are observable and measurable" (Baujard 2017, 88). According to Paretian theory, interpersonal comparison based on cardinal utilities are not operational because they have no empirical significance. In Chapter 1 of this dissertation, the topic of interpersonal comparisons of utility is useful to analyze the value judgments of economists and economic agents. We focus on Harsanyi's attempt to justify interpersonal comparisons of utility without making value judgments, based on his impartial observer theorem.

The second form of normative economics is the so-called *new welfare economics*, which sets up a clear separation between the optimality conditions in the social sphere and the optimality conditions for the functioning of the market. It is important to say that the breaking point of this second form of normative economics with respect to the first lies in the impossibility of making interpersonal comparisons of utility and in considering ordinal utility information. In order to study the optimality conditions, the economists of the *new welfare economics* focus on the Pareto principle, leaving aside any evaluation of income distribution (Baujard 2017).[11]

Two schools of the new welfare economics are identified by Suzumura (2021). On the one hand, there is the Bergson-Samuelson school. Abram Bergson (1938) introduced the concept of a social welfare function and Paul Samuelson (1947; 1981) develops and disseminates this concept.[12] On the other hand, there is the Kaldor-Hicks school who proposes compensation

[10] Robbins says elsewhere in the same book that
"the theory of exchange […] does not assume that, at any point, it is necessary to compare the satisfaction which I get from the spending of 6d. on bread with the satisfaction which the Baker gets by receiving it […] It is a comparison which necessarily falls outside the scope of any positive science […] It involves an element of conventional valuation. Hence it is essentially normative. It has no place in pure science" (Robbins 1935 [1984], 138–39).

[11] The Pareto principle establishes that if individuals are unanimous in ordering a pair of options, this ordering is reflected in the social preference. Pareto optimality implies that it is not possible to improve the situation of any individual in society without deteriorating the situation of at least another one. See Chapter 2 of Sen (1970 [2017]) for a formal definition of the Pareto criterion, as well as his discussion of the strong and weak forms of the Pareto criterion.

[12] See Backhouse (2021) and Suzumura (2021), who discuss the differing views of Bergson and Samuelson regarding the nature and origin of the Bergson-Samuelson social welfare function.

principles according to which potential transfers would allow Pareto optimal comparisons.[13] In brief, the new welfare economics seeks to deduce from individual preferences the best collective decision, without considering interpersonal comparisons of utility, which were considered non-scientific.

Arrow's *Social Choice and Individual Values* (1951) creates a major setback for the new welfare economics by stating that it is not systematically possible to obtain a social welfare function based on individual preferences under a list of normative constraints. This leads to the emergence of two more forms of normative economics: social choice theory and public economics. On the one hand, social choice theory "investigates the various abstract methods of evaluating social states." On the other hand, public economics is dedicated to studying "market optimality and policy corrections" (Mongin 2002, 147). A fifth form of normative economics would be considered according to Mongin (2002), which focuses on the study of theories of justice and equity.

What is a Value Judgment?

Let us to present the definition of a value judgment proposed by Fisher (1913): "A judgment of value […] is a thought about an object, a thought in which value for some subject is attributed to the object, or predicated of it" (Fisher 1913, 626). For instance, in "A is good," the value for the subject "A" is given by the predicate "good." Another way to analyze what a value judgment is, is to try and identify what a value judgment is not. Grünberg (1978) proposes three ways to identify what a value judgment is not: (1) "The value judgment is not a judgment about value." It means that in the value judgment "the predicate of the judgment is a value discernment," whereas in the judgment about value "the subject of the discernment is a value ("Beauty may be contemplated" […]"); (2) "The value judgments is not a discernment which asserts a value in the way the following sentences do: 'This flower is worth $2,' or 'This man is respected by his friends.'" It means that the person who makes a value judgment is not simply making a statement, nor neutrally confirming the value given by other people; (3) "The value judgments

[13] The compensation principles are introduced by Nicholas Kaldor (1939) and John Hicks (1940). A well-known further elaboration is made by Tibor Scitovsky (1941). Assuming the social situations x, y, z. Kaldor's criterion establishes x is preferable to y, if and only if, starting from x, it is possible to reach a state z by transfers, such that the potential situation z is Pareto-superior to y. Hicks's criterion establishes x is preferable to y, if and only if, starting from y, it is not possible to reach a state z by transfers, such that the potential situation z is Pareto-superior to x.

is not a simple statement about a preference." It means that sentences such as "I like Z," I don't like Z," "I prefer X to Y" are not value judgments because they state a preference (a mere fact), or an emotional content only. (Grünberg 1978, 127).

In general, in the philosophy of science, the linguistic characteristics of value judgments are widely studied. This has not been however the case in economics. Economists have rather concentrated on studying the distinction between positive and normative economics in terms of fact and values. Economists such as Backhouse (2005), Blaug (1992; 1998), Hutchison (1964), Röpke 1941 [2015], among others, have focused on the study of the role of value judgments in economics. In general, these authors accept value judgments of a methodological nature, that is, value judgments that have an influence on the choice of a study topic, as well as the approaches and methods to carry out research. However, the differences between the points of view of these authors arise when they analyze the impact that value judgments have when accepting or rejecting a hypothesis or theory, when studying economic concepts, or when giving practical advice on economic policy.

Mongin (2006b) notes that economists have neglected linguistic characteristics when studying value judgments. "The literature of economics and the social sciences, from Weber through Robbins and Myrdal and up to Sen and Arrow, emphasizes value judgments at the expense of value sentences, statements and utterances"(Mongin 2006b, 262). He analyzes the nature of predicates to determine what kind of judgment the economist makes. He thus examines thick predicates, which are simultaneously evaluative and descriptive. This mirrors the same idea as Putnam's thick concepts, suggesting that judgments of values and judgments of fact may be entangled. For this reason, in the first chapter of this dissertation we study the way in which Mongin (2006b) distinguishes value judgments from other judgments, which leads him to propose a classification of the degree of neutrality of economists when making value judgments.

It is important to be clear about two types of judgments. First, judgments of facts which refer to descriptive statements, for instance "Fiscal policy is an economic policy, so is monetary policy." Second, judgments of values, which refer to evaluative statements, for example "Expansionary fiscal policy is a good economic policy." Mongin asserts that evaluative statements cannot have the same logical form as prescriptive statements. For this reason, he does not consider prescriptive statements ("The government should implement expansionary

fiscal policy") as judgments of values.[14] Normative economics includes evaluative and prescriptive statements while positive economics includes descriptive statements. However, it cannot be ruled out that judgments of value and judgments of fact are entangled, so that it will be difficult to identify them in normative economics or positive economics. It implies that value judgments are also present in the analysis of positive economics.[15] However, we focus on elucidating the role of value judgments in normative economics as our primary interest is here to reappraise normative economics as a science.

Outline of the Dissertation

In order to determine the role played by value judgments in the study of normative economics, we proceed as follows. In Chapter 1, we address two questions: (1) To what extent are economists willing to make value judgments for themselves when studying normative economics? (2) How do economists take into account the value judgments of economic agents in the study of normative economics? We claim that the distinction between the value judgments of economists and those of economic agents has not been explicitly established in welfare economics. Welfare economics has focused on the debate around the scientific status of normative economics, so that the relevance of value judgments of economists has been at the center of discussion. As a result, this distinction is necessary as it leads economists to consider not only their own value judgments but also those of economic agents within normative economics. It is essential to have a theoretical framework to make this distinction. For this purpose, we focus on two approaches that can be useful: Sen's (1967) classification of value judgments and Mongin's (2006b) theses about value neutrality in economics. Finally, we take the subject of interpersonal comparisons of utility as a case study in order to show the importance of this distinction, more specifically the analysis of Harsanyi's impartial observer theorem. This chapter will allow us to justify why economists can make value judgments for themselves in the study of normative economics and why it is important to have a theoretical framework to analyze the distinction between the value judgments of economists and those of economic agents.

[14] "Evaluations have a predicative form – i.e., they are logically formalized in terms of n-place predicates – but prescriptions have a modal form – i.e., they are logically formalized in terms of modal operator ('It is obligatory that') acting or sentences." (Mongin 2006b, 267).
[15] See Małecka (2021) and Reiss (2017) who take into account the literature of the philosophy of science to study how value judgments are present in the analysis of positive economics.

In Chapter 2, we address the following question: (3) How do normative economists study and characterize a value (freedom) in relation with the value judgments of economic agents (as represented by their preferences)? We analyze the role of freedom as a normative criterion for evaluating alternative states of affairs. We focus on analyzing the freedom of choice that an individual has over an opportunity set at a given moment. It refers to "the opportunity that we have to live the way we would like, do the things we would choose to do, achieve the things we would prefer to achieve" (Sen 1990, 471). Our purpose is to determine how the opportunity or preference-based approaches use the tools of social choice theory to understand the conception of freedom in economics. We study preference-based approaches because they have been strongly challenged and we think that their contribution to the conception and value of freedom has to be highlighted. The importance of Sen's contribution (1990; 1991; 1993b) to the development of the preference-based approaches is examined. However, since his approach only considers actual preferences, we also study the importance of considering the uncertainty of individuals regarding future preferences. This has been studied by authors such as Arrow (1995), Foster (1993), Kreps (1979), Puppe (1996), among others. Both the actual preference approach and the uncertainty preference approach were criticized by Jones and Sugden (1982) and Pattanaik and Xu (1998) and for not considering that "freedom of choice is important for its own sake, irrespective of any consequences it may entail for the chooser" (Gravel, Laslier, and Trannoy 1998, 78). The analysis of the different preference-based approaches to measure freedom of choice allows us to establish the influence of social choice theory in understanding the conception of freedom in economics.

From the study of the single value of freedom, we go on to study in the last chapter of this dissertation (Chapter 3) the role of value judgments in the elaboration and interpretation of Arrow's impossibility theorem. We address the following question: (4) To what extent are value judgments important in the elaboration and interpretation of normative economic theories? We at first try to understand how Arrow's own values and his cultural context influenced his methodological choice. In order to identify the methodological value judgments made by Arrow to obtain his impossibility result, we examine the storytelling proposed by Arrow regarding the elaboration of his impossibility theorem approximately thirty years after the first publication of *Social Choice and Individual Values*. We focus on interviews of his colleagues with Arrow and on scientific articles where Arrow explains the origins of his impossibility theorem. We then analyze Arrow's impossibility theorem under the prism of Mongin (1999; 2002; 2006b). We consider three reasons why Mongin's contributions shed light on Arrow's position vis-à-vis the

role of value judgments in his impossibility theorem. First, Mongin (1999) discusses the importance of value judgments in Arrow's impossibility theorem and the impact this has in normative economics. Second, Mongin (2002) studies the importance of Arrow's impossibility theorem in the progress of normative economics. Third, Mongin (2006b) proposes a value neutrality classification that allows us to determine what position Arrow assumed regarding the role of value judgments in his impossibility theorem. This analysis allows us to study the value judgments that led Arrow to concentrate on loosening either the unrestricted domain condition or the independence of irrelevant alternatives condition of his impossibility theorem.

Chapter 1

The Value Judgments of Economists and of Economic Agents: The Case for a Distinction

---

Abstract

The distinction between the value judgments of economists and those of economic agents is not clear in the literature on welfare economics. In this chapter, we provide a framework to justify that the value judgments of economists and those of economic agents have to be properly identified. First, we propose a comparison between two renowned authors in the discussion of the role of value judgments in social sciences: Max Weber and Gunnar Myrdal. Second, we study two approaches that can be useful in order to make the distinction: Sen's classification of value judgments and Mongin's theses about value neutrality in economics. Finally, we use the subject of interpersonal comparisons of utility to discuss this distinction, more specifically the analysis of Harsanyi's impartial observer theorem. This allow us to show that it is not possible to 'objectify' subjective attitudes of economic agents, and that economists cannot avoid value judgments when they recommend an action for the benefit of the society.

1.1 Introduction

The distinction between the value judgments of economists and those of economic agents is not clear in the literature on welfare economics. We suggest that the importance of making this distinction lies in determining whether economists may make value judgments in their professional work and how value judgments may be crucial to justify the economic agents' preferences. In essence, we analyze this distinction, not in a separated but a related way. This means that we do not limit ourselves to making the distinction, but we seek to show its meaning and its usefulness in normative economics.

The degree of neutrality that economists have when they are asked for advice regarding economic policy has been a crucial debate in the field of welfare economics. On the one hand, the upholders of the idea that economists are analytical technicians that study an economic problem rationally without entangling their value judgments refer to the well-known notion of axiological neutrality proposed by Weber (1919 [1965]). It establishes the objectivity of knowledge that a social sciences scientist has to put into practice when he does research. This objectivity consists of the fact that the economist is conscious that values are present when deciding the subject of research and the direction it is going to take. In this way, the economist can take a position regarding his values. On the other hand, the upholders of the idea that economists may study economic problems based on not only facts but also on value judgments refer to Myrdal's proposal: "The interests can never be purely scientific. They are choices, the products of our valuations. Without valuations we have no interest, no sense of relevance or of significant and, consequently, no object" (Myrdal 1958, 51).

Welfare economics has been at the center of this debate because it provides a theoretical framework that allows assessing the consequences of individual actions and public decisions on social states (Baujard 2017). Behind individuals' actions, there are the value judgments of individuals that I shall now call "economic agents." Counseling economic agents is economists' concern, and this counseling is meaningful if it is based on values of economic agents (Bergson 1954). A subject that relates the value judgments of economists to those of economic agents is

interpersonal comparisons of utility (ICU's). Accepting or rejecting ICU's has been crucial in the history of welfare economics and therefore in normative economics' progress.[16] The rejection of ICU's has been a key element in the transition from the "old" to the "new" welfare economics. However, ICU's will be reintroduced in welfare economics in order to avoid Arrow's impossibility theorem (Backhouse, Baujard, and Nishizawa 2021, 5–6) . In the "old" welfare economics, ICU's can play an essential role in determining economic policies to improve people's well-being. The way ICU's are made reveals a normative conception. In the "new" welfare economics, ICU's are rejected because these comparisons are based on value judgments and not on empirical facts and also because the economist cannot match them with possible experience (Mongin 1999).

Although these two positions on interpersonal comparisons of utility seem irreconcilable, Harsanyi (1955; 1977) argues that it is possible to make ICU's, without resorting to value judgments. The analysis of Harsanyi's approach to ICU's allows one to study how an economist can claim that ICU's are not based on ethical or political value judgments. This is because Harsanyi assumes that individual preferences are based on impersonal attitudes (extended preferences) guaranteed by the existence of an impartial observer. We use Harsanyi's attempt to 'objectify' ICU's to analyze the distinction between the value judgments of economists and those of economic agents in welfare economics. This distinction has not been explicitly established in welfare economics, which has concentrated on the debate on economics as a neutral social science. As a result, economists have to step back and consider not only their own value judgments but also those of economic agents when they study economic phenomena. So, my contribution is to bring this distinction to light and propose a theoretical framework that allows it to be better studied. This theoretical framework will be useful in examining the nature of the value judgments of economists and economic agents, which will allow us to clarify the role of value judgments in normative economics.

All this considered, this chapter has five sections. The first section presents a comparison between Myrdal and Weber, which is focused on the role of value judgments in social sciences, and a comparison between Mongin and Sen, which is focused on the role of value judgments

[16] Mongin(2006a) studies the progress in normative economics. He proposes a provisional progress definition, which will be his main tool for analyzing what he calls the intertheoretic progress in normative economics. Thus, Mongin suggests the following definition of progress:
"A shift from theory T to theory T' is progressive if: (1) T' provides a solution to at least one unresolved problem of T; (2) T' provides a solution to the main problems that T had already addressed and resolved in its own way; (3) T' raises new problems and manages to solve at least one of them; (4) T does not satisfy the previous conditions with respect to T'" (Mongin 2006a, 25).

in economics. The second section analyzes the value judgments of economic agents in welfare economics. The third section studies the value judgments of economists in their professional work. The fourth section presents a case study based on interpersonal comparisons of utility with the aim of showing the importance of the distinction between the role of values of economists and that of economic agents in economics; and the fifth section concludes.

1.2 At the Crossroads of Dealing with the Role of Values in Social Sciences and in Economics

The contributions of Weber and Myrdal are essential in the discussion of the role of value judgments in social sciences. Weber is one of the most important modern sociologists. His work refers not only to sociology but also to social sciences in general. His thinking is important in establishing concepts essential in economic and sociological analysis, such as axiological neutrality. Myrdal is an economist concerned to show the complexity of social phenomena and the gulf between social research and natural research. He is one of the most important upholders of the importance of value premises in social sciences.

These two authors have often been opposed in the literature because they present different ways of conceiving values in social sciences. Their opposite conceptions may be useful to understand the debate about the difficulty for social scientists to take values into account. For this reason, their ideas have been considered by economists (cf. Backhouse (2005); Blaug (1992, 1998); Boumans and Davis (2010); Mongin (2006b)) in order to study methodological value judgments in economics, the value-ladenness of economic concepts and explanations, and the way explanations in economics need to take into account people's value judgments.

Once analyzed how Myrdal and Weber consider value judgments in the social sciences, we focus on laying the ground of the theoretical framework to make the distinction between the value judgments of economists and those of economic agents. To do this we decide to compare the framework of Mongin (2006b) and that of Sen (1967) of the role of value judgments in normative economics. Mongin is an economist that has focused simultaneously on economic theories and on the philosophy of economics in social sciences. At the end of 20th century, he has contributed to the international development of the philosophy of economics. His typology of the degree of neutrality of economists in their professional work is essential in order to study the different alternatives to solve the value neutrality problem. Sen is an economist and a

philosopher who has made great contributions to welfare economics, to social choice theory and to problems associated with poverty. He is considered as one of the greatest upholders of the importance of values in normative economics.

In brief, we first present the main ideas defended by Weber (1904 [1949]) and Myrdal (1958) regarding the role of value judgments in social sciences. Then we proceed to compare them with the aim to examine if they identify the role of value judgments of social scientists and of individuals in social sciences. We secondly present the theoretical grounds of Mongin (2006b) and Sen's (1967) frameworks, which will allow us to compare the tools they use when studying value judgments in normative economics.

1.2.1 Myrdal & Weber: The Role of Value Judgments in Social Sciences

1.2.1.1 Myrdal's framework: Getting Rid of Value Premises is not an Option

We focus on the analysis made by Myrdal (1958) in his book *Value in Social Theory: A Selection of Essays on Methodology*, where he proposes a relevant background to his conception of values in social sciences by way of the *Negro Problem* in the United States. We are especially interested in the difference between valuations and beliefs, the biases that can be present in social sciences and the role of valuations to correct them, and the value premises in research. All these points seem to us important in order to understand the position defended by Myrdal on the role of values in economics.

First of all, a person's valuations refer to her "ideas about how it ought to be, or ought to has been", whereas a person's beliefs refer to her "ideas about how reality actually is or was" (Myrdal 1958, 71). Concerning valuations, they can be part of different dimensions of a person's moral personality. It means that if a person considers the valuations belonging to one aspect of her morality, the valuations belonging to other spheres of her morality are not considered at the same time. Myrdal tries to show that it is essential that social studies consider a person's conflicting valuations, because this will help to understand the case study of the *Negro Problem* better. The problem is that some white people deny Negroes the same civic rights as themselves, while in another sphere of their moral personality, they have valuations in

favor of liberty, equality, justice and fair opportunity for everybody (which are the values of the American society) (Myrdal 1958, 59).[17]

Concerning beliefs, they serve as the basis for the construction of rational hierarchies of valuations, which allow the shaping of a person's opinions. If valuations are conflicting, as is often the case, "beliefs serve the function of bridging illogicalities" (Myrdal 1958, 76). Myrdal considers that valuations and beliefs can form opinions, and that beliefs and values can influence each other. The *Negro problem* helps to explain the role of beliefs. There are different beliefs that are surrounding the Negro and his status in American society. If one of these beliefs is that the Negro belongs to a lower race (evidently a false belief), some white people may adopt this belief in order to justify their way of life. But, if there is a campaign that shows some white people are wrong in how their way of life ought to be, they will be pressured to change their valuations in some way. In other words, if the campaign succeeds in changing the popular belief about the racial inferiority of a part of the population, then white people have to readjust the hierarchy of their valuations, and possibly their behavior. In a nutshell, through this example based on Myrdal but slightly modified, we observe how valuations and beliefs influence each other in reality.

An essential feature of Myrdal's distinction is the identification of a set of valuations that a person has according to the different dimensions of her moral personality. This proposition could be considered as a primitive idea of the plural identities of a person developed by Sen (2006) and Binder (2019). In essence, a person can not only be a Venezuelan citizen, a student and an only child, but also a cinephile and a tennis fan, where each of these plural identity dimensions comes with a set of values, objectives and practices.

A second aspect that we consider as important is the scientific biases that Myrdal identified in his case study (the Negro problem in the US), which could be presented in other cases in social sciences as well. Among the biases found, we have first the scale of friendliness (the degree of sympathy that a scientist shows for the feature studied). Second, the scale of radicalism-conservatism (it refers to the way the political and economic thoughts of scientists may affect their research). Third, the scale of optimism-pessimism (it denotes the influence that encouraging and discouraging statements may have on scientists). Fourth, the scale of isolation-

[17] We use the word "Negro" to be consistent with the vocabulary used by Myrdal in his book *An American Dilemma* (1944). This word is never used with a negative connotation in this dissertation.

integration (it refers to the degree to which scientists may establish a relationship between the object of study and the economic, political, social and even judicial context). Fifth, the scale of scientific integrity (the ability of scientists to study unpopular subjects from which they might obtain controversial results).

Despite these biases, Myrdal does not discard valuations in social sciences, because he affirms that a social science is a political science and that problems must be faced based on not only facts but also on valuations. In the following quotation Myrdal's position is clear:

> The valuations will, when driven underground, hinder observation and inference from becoming truly objective. This can be avoided only by making the valuations explicit. There is no other device for excluding biases in the social sciences than to face the valuations and to introduce them as explicitly stated, specific, and sufficiently concretized, value premises (Myrdal 1958, 132).

In this way, Myrdal proposes that it is not by hiding valuations behind so-called facts that the biases explained above can be overcome. For this purpose, social scientists choose their value premises according to the relevance of the case of study, and that allows them to be introduced rationally into social sciences. In other words, not only the observation of the facts of a given situation can explain whether this situation is suitable or desirable, but also the value premises considered relevant for the scientist.

If the value premises have to be taken into account explicitly to avoid biases in social sciences, then they have to meet some criteria in the research activity: be explicitly stated, be specific and concrete, be intentionally selected, have a hypothetical character, constitute a set of alternatives hypotheses, have a relevance, be significant, be feasible and be consistent. Nevertheless, the selection of value premises is not an easy task given that a person can have different types of valuations. Those valuations can be based on correct or incorrect beliefs (beliefs and values can influence each other), and it is difficult to deal with many relevant sets of valuations.

In order to avoid these difficulties, Myrdal proposes an "instrumental norm", which consists of two stages: in the first stage the social scientist takes one single set of relevant and significant value premises in order to make an initial analysis of an existing situation, and in the second stage other significant sets of value premises are introduced with the aim of allowing judgments in terms of alternative valuations and policies.

1.2.1.2 Weber's framework: Escaping from the Researcher's Subjectivity

In the *Methodology of the Social Sciences*, Weber (1904 [1949]) distinguishes two important definitions, that of value judgments and that of relevance to values. Value judgments are "practical evaluations of the unsatisfactory or satisfactory character of phenomena subject to our influence" (Weber 1904 [1949], 1). By contrast, relevance to values denotes "the philosophical interpretation of that specifically scientific 'interest' which determines the selection of a given subject-matter and the problems of an empirical analysis" (Weber 1904 [1949], 22).[18]

Weber establishes a clear separation between empirical facts and practical evaluations made by the researcher. In this way, it is not the researcher's role to determine whether the empirical facts are satisfactory or not. However, the scientific importance of evaluations is grounded on two reasons:

> […] It also has high scientific importance: (1) for purposes of an empirical causal analysis which attempts to establish the really decisive motives of human actions, and (2) for the communication of really divergent evaluations when one is discussing with a person who really or apparently has different evaluations from one's self. The real significance of a discussion of evaluations lies in its contribution to the understanding of what one's opponent (or one's self) really means (Weber 1904 [1949], 14).

As a result, value judgments are not entirely ruled out from scientific debates. To defend this proposition, Weber assumes that human behavior is concerned with the categories *ends* and *means*. The determined action of a person is contingent upon reaching an *end* through specific available *means*. In the decision-making process, a person has to choose according to her values and conceptions about the world, and the role of the social scientist is to provide a person with the possibility of weighing and comparing the desirable and undesirable consequences of her action. In essence, "the scientific treatment of value judgments may not only understand and empathically analyze the desired ends and the ideals which underlie them, it can also 'judge' them critically" (Weber 1904 [1949], 54).[19]

Value judgments will however not be decisive in the choice of a research topic in social sciences, but rather in the value-relevance of the research topic treated. In other words, the

[18] We base our analysis on two of the three essays of the translated edition of *The Methodology of Social Science*: "The Meaning of 'Ethical Neutrality' in Sociology and Economics; and "'Objectivity' in Social Science and Social Policy."

[19] "An empirical science cannot tell anyone what he should do but rather what he can do (and under certain circumstances) what he wishes to do" (Weber 1904 [1949], 54).

social scientist will study the research topic that seems important according to the cultural context and her own values. As a result, without the social scientist's interest in the importance of cultural facts, it is not possible to have the principle of choice of a phenomenon to be studied. In the following quotation this point is quite clear:

> When we require from the historian and social research worker as an elementary presupposition that they distinguish the important from the trivial and that he should have the necessary "point of view" for this distinction, we mean that they must understand how to relate the events of the real world consciously or unconsciously to universal "cultural values" and to select out those relationships which are significant for us (Weber 1904 [1949], 82).

In relationship with the value-relevance that a phenomenon can have for its study, Weber proposes the ideal-type definition. The ideal-type is precisely a logical construction that "has no connection at all with value judgments" (Weber 1904 [1949], 98). This reasonable construction serves as a guide for the formulation of hypotheses to study a phenomenon. In essence, the ideal-type consists of picking out the relations of interest and isolate them from perturbations that interfere with their mechanism in the real world. For this reason, this analytical construct (*Gedankenbild*) cannot be directly applied, but it can be a point of reference in order to answer questions about a social phenomenon.

When social scientists decide to study a phenomenon, they have to choose the most important elements according to cultural facts. This process will allow them to construct the ideal-type, which will be the point of comparison with the real world. This comparison will allow them to describe the divergences and similarities between the utopian construction and the reality "with the most unambiguously intelligible concepts, and to understand them causally" (Weber 1904 [1949], 43). The value-relevance and the ideal-type are both essential to maintain the researcher's objectivity in the sphere of social sciences.

To explain the concept of ideal-type effectively, we can use the framework developed by Morgan (2012), who proposes four ways to delineate economic models. One of the ways of modeling considered by Morgan is idealization. She explains this process through different portraits of the economic man along with the history of economic thought and she demonstrates how his role has changed depending on the relations of interest that economists isolated to define him. "Malthus portrayed him as driven by physical appetites, Mill as a wealth seeker. Jevons changed him into a man seeking to maximize pleasure or utility from consumption,

while Menger presented him as satisfying needs though sensible choices" (Morgan 2012, 164).[20]

We note that each portrait of the economic man is defined according to the economic principles of economists (the fact of belonging to a school of economic thought), and the cultural context in which these economists were. Furthermore, these portraits allow us to "reduce the complexity of dealing with all human feelings and emotions and actions that flow from them and, at the same time to focus the attention on the explicitly economic aspects of man's behavior" (Morgan 2012, 165).

1.2.1.3 Comparison between Myrdal's framework and Weber's framework

Even if it is evident that Myrdal and Weber are poles apart concerning their conceptions about the role of valuations in social sciences, there seems to be a common feature that deserves to be discussed. We noted how the cultural context plays an important role for the two authors when the social scientist decides to study a phenomenon. Myrdal clearly said, "the value premises should be selected by the criterion of relevance and significance to the culture under study" (Myrdal 1958, 134); and Weber mentioned the importance of the cultural context in his value-relevance conception.

The cultural context is used by each author according to what they want to show about the degree of objectivity of the social scientist. On the one hand, Myrdal is convinced that full objectivity can never be reached, because the social scientist is part of the culture that he tries to study, and that generates scientific bias. On the other hand, Weber defends the objectivity of the social scientist, which is present when the researcher chooses a phenomenon to study based on the cultural context (value-relevance) and when he analyses this phenomenon through the ideal-type.

It is important to note that both authors try to avoid biases in the study of social phenomena, but in a different way. Myrdal argues that value premises do not have to be concealed behind facts, because the biases in social sciences are given by not making explicit the valuations of the scientist. The following quotation shows the point defended by Myrdal:

[20] See Chapter 4. Character Making: Ideal Types, Idealization, and the Art of Caricature to deepen in the subject.

> There is a common belief that the type of practical research which involves rational planning (what we have ventured to call 'social engineering') is likely to be emotional. This is a mistake. If the value premises are sufficiently, fully, and rationally introduced, the planning of induced social change is no more emotional by itself than the planning of a bridge or the taking of a census […] Emotion and irrationality in science, on the contrary, acquire their high potency precisely when valuations are kept suppressed or remain concealed in the social so-called 'facts' (Myrdal 1958, 133).

A critical limit of this approach is the way value judgments are selected and introduced rationally into scientific research. Even Myrdal faces this problem when he studies the *Negro problem*. For this purpose, he proposes the "instrumental norm", which is represented by what Myrdal call the *American Creed* ("ideals of the dignity of the individual, fundamental equality, and certain inalienable rights to freedom, justice and fair opportunity") (Myrdal 1958, 65). In essence, Myrdal chooses these ideals to analyze the *Negro problem* in the United States because they are part of the American culture. On the contrary, to avoid biases in social sciences, Weber conceives the ideal-type construction, which allows him to keep value-judgments out of scientific analysis. So, the cultural context will be essential to define the research object, but not to select the most relevant value premises.

With regards to Weber's framework, we observe that he identifies the role of value judgments of social scientists and those of individuals in social science. In the case of an individual, his value judgments will be essential for making choices, and he chooses "from among the values involved according to his own conscience and his personal view of the world" (Weber 1904 [1949], 53). Concerning social scientists, they help an individual "in becoming aware of the ultimate standards of value which he does not make explicit to himself or, which he must presuppose to be logical" (Weber 1904 [1949], 54). As a result, value judgments are essential in the individual sphere at the moment of making choices. In the scientific sphere, the role of value judgments is limited to understanding them and also judging them logically.

Like Weber, Myrdal identifies the role of the value judgments of social scientists and individuals in social sciences. However, the importance that Myrdal gives to valuations in the scientific sphere is different from Weber. Social scientists have to select and introduce value premises in their analysis in order to avoid concealing valuations behind facts. "Practical conclusions may thus be reached by rational inferences from the data and the value premises. Only in this way does social engineering, as an advanced branch of social research, become a rational discipline under full scientific control" (Myrdal 1958, 132). Concerning individuals, valuations and beliefs are essential for determining their behavior. According to Myrdal, even

the ordinary man tries however to be 'objective' and gives 'reasons' for his desires. This means that individuals present their valuations as rational beliefs about reality. In brief, to choose the value premises in order to study a social phenomenon, the social scientist has to consider the set of valuations and beliefs of individuals.

On balance we can say that both authors use the cultural context in their framework to obtain essential information (relevant value premises or relevant subject-matter), and both authors make the distinction between the role of value judgments of social scientists and that of individuals in social sciences. The fact that both authors make this distinction, shows us that value judgments are perceived differently depending on whether someone takes the place of a scientist or that of an individual.

In the next sub-section, we focus on laying the foundations of Mongin (2006b) and Sen's (1967) frameworks, which will be essential to establish a distinction between the value judgments of economists and those of economic agents. On the one hand, we will study the importance that Mongin gives to the linguistic character in the study of value judgments, which allows him to propose a classification of the degree of neutrality of the economist. On the other hand, we will analyze Sen's attempt to study the nature of value judgments and its usefulness to examine the value judgments of economic agents.

1.2.2 Mongin & Sen: Evaluations, Prescriptions and Obligations

1.2.2.1 Mongin's ideas about Economic Evaluations, Prescriptions and Obligations

Regarding economic evaluations, Mongin (2006b) starts by separating judgments of fact and judgments of value. When a judgment of value is made, he distinguishes four characteristics: the linguistic entity which is represented by the *sentence* itself; the sentence meaning which is identified as the *statement;* the material act performed by the speaker or writer which is the *utterance or inscription*; and the *judgment* being made. One of the reasons to consider these four items in an economic evaluation according to Mongin, is because it allows one to identify when economists are reciting a sentence or making a judgment.

To determine if economists make judgments of facts or judgments of value, Mongin proposes a two-stage analysis. At stage 1, a diagnosis of the statement expressed by the sentence is made. The idea is to determine whether a non-analytical statement can be considered as either logically

evaluative, or logically factual or logically neutral.[21] Each of these logical classifications are related to the nature of the predicates in the statement. There are three sub cases: (1) if at least one evaluative predicate is used in the statement, it is a logically evaluative statement; (2) if factual or logical predicates are used in the statement, it is a logically factual statement; and (3) the statement is logically neutral in residual cases.

Mongin states that the evaluative or factual predicates must be used in the statement and not just mentioned. If they are merely mentioned in the statement, then the statement does not count as being logically evaluative or factual. In the sentence '"Rational" is an eight-letter word,' the predicate 'eight-letter word' is used for stating what 'rational' is, and this is then a logically factual statement. On the other hand, in the sentence 'X is a good economic policy', the predicate 'good' is used for evaluating the economic policy, and this is then a logically evaluative statement (Mongin 2006b, 263–64).

Stage 2 restricts the analysis to those statements that are regarded to be either logically evaluative or logically factual. With this in mind, "the economist makes a judgment of value if the statement is logically evaluative and the economist sincerely asserts it; he makes a judgment of fact if the statement is logically factual and he sincerely asserts it; and he makes no judgment at all if he does not make a sincere assertion" (Mongin 2006b, 263).

Mongin analyses stage 1 through a distinction between thin and thick predicates. The thinness of a predicate is characterized by being exclusively evaluative. By contrast, the thickness of a predicate is characterized by being evaluative and descriptive at the same time. For instance, when Joshua says that Anne is good, he seems to be evaluating her without giving much description. In other words, Joshua only evaluates but does not describe Anne because he does not explain in what Anne is good.[22] On the contrary, in the case of a thick predicate, when Joshua says that Anne is honest, he appears to be doing two things: evaluating her positively and describing her as someone truthful and sincere.

The importance of this distinction between predicates is crucial in order to identify the evaluative side of a predicate of economics. Mongin makes however clear in the following quotation the difficulty of finding a criterion to do it:

[21] "While analytical statements are automatically endowed with a truth value, non-analytical statements may or may not have one" (Mongin 2006b, 284). Mongin does not consider analytical statements in his analysis because those statements can be declared true or false by virtue of the meaning of their constituent terms (Ibid, 264).
[22] "A necessary condition for 'good' to be descriptive is that 'X' in the sentence must be a functional word, i.e., a word the definition of which explains what X is for" (Hare, 1957, comment by Mongin 2006b).

> The basic objective is to become able to recognize whether or not a predicate of economics has an evaluative side. While admitting that I do not provide a criterion, I stress that philosophers have done no better. What their work offers is a list of necessary conditions for a statement to count as an evaluation that have the extra property of not being also necessary for the statement to count as a statement of fact (Mongin 2006b, 271).

In addition to this distinction between thick and thin predicates, Mongin identifies ambiguous predicates,[23] which can sometimes be interpreted either in a factual or in an evaluative manner. He emphasizes that a thick predicate cannot be confused with an ambiguous predicate, because the former is descriptive and evaluative at the same time (two sides of the same coin), while the latter can be alternatively factual or evaluative.

Another important point highlighted by Mongin is the difference between evaluations, prescriptions and obligations. Evaluations have a predicative form, which means that they are logically formalized in terms of n-place predicates; while prescriptions have a modal form, which means that they are logically formalized in terms of a modal operator acting on sentences. For instance, if we compare the following statements: (1) 'X is a better economic policy than Y' and (2) 'Society should implement economic policy X rather than Y'. The former expresses evaluation and the latter expresses a prescription. If we analyze these statements, we could say that (1) can bring about (2) (Mongin 2006b, 266–67).

Even if obligations have the same general logical structure than prescriptions, obligations are often less specific than prescriptions. It means that obligations may be vague with regards to the relevant actions to be taken, while prescriptions must specify them. If we compare statement (3) 'Society should implement economic policy X' with (2) in the previous paragraph, (3) expresses an obligation because the statement states a result to be achieved without specifying a context. In short, Mongin considers that the difference between these three terms (evaluation, obligations and prescriptions) is significant for economists, because "it is important to be clear about whether economic policy is just a valuable policy to follow, or it tells the policy-makers how to shape their actions in the broad, or it sets concrete steps for them to take" (Mongin 2001b, 11-2).

It should be pointed out that this framework established by Mongin is useful for his analysis of the value neutrality problem in economics, which he studies using four theses: Strong Neutrality

[23] An example of an ambiguous predicate is *efficient*: "In economics, the predicate efficient ranges on a wide array of meanings, some of them clearly factual, other frankly evaluative" (Mongin 2006b, 266).

(SN), Weak Neutrality (WN), Strong Non-Neutrality (SNN) and Weak Non-Neutrality (WNN). In essence, he uses the two-stage approach that we explained above as the basis of his analysis, to elucidate why so many economic statements cannot fit in well with what he calls the authoritative criterion. This criterion is based on the positive-normative distinction, and it can be defined as follows:[24]

> […] A statement made by an economist counts as normative if it is paired with a value judgment made by that economist (not somebody else), and as positive otherwise (in particular, it may bear on somebody else's value judgments) (Mongin 2006b, 259).

What Mongin does is to question the exhaustive and exclusive character of the authoritative criterion. He demonstrates that economic statements cannot always be qualified as either positive or normative (not exhaustive), and that economic statements can be qualified as positive and normative at the same time (not exclusive). Besides, Mongin claims that this criterion can sometimes involve a counterintuitive division between the positive and the normative.[25] All these weaknesses of the authoritative criterion that Mongin corroborates through his analysis lead him to the following conclusion about the role of value judgments of economists in their professional work:

> Economists make and are justified in making a large number of value judgments; that some of these are easy to pinpoint, but others are not; and that judgments of fact and of value often turn out to be inseparable from each other (Mongin 2006b, 282).

1.2.2.2 Sen's ideas about Economic Evaluations and Prescriptions

Sen (1967) proposes a classification of prescriptive judgments, which allows him to study the nature of value judgments. He begins by making a difference between what he calls a "purely prescriptive judgment" and an "evaluative judgment". Sen relies on Hare's (1952) distinction between a "prescriptive" term, which has prescriptive meaning regardless of whether it has descriptive meaning or not, and an "evaluative" term, which has both types of meaning. Hare's distinction is essential for Sen because it allows him to establish his first criterion of value

[24] The four theses will be explained in detail in sub-section 1.4.2 of this chapter.
[25] We will analyze the examples that Mongin gives to justify his criticism of the authoritative criterion in sub-section 1.4.2 of this chapter, since in this sub-section we will examine how Mongin used the tools exposed in this sub-section (1.2.2.1) to establish his classification of the degree of neutrality of the economists.

judgments: purely prescriptive and evaluative judgments (Sen 1967, 46). Sen defines two other criteria to classify value judgments: compulsive and non-compulsive judgments, and basic and non-basic judgments. The former criterion concerns judgments based on imperatives, which can be made either without considering the rest of someone's value system or considering someone's conceivable value judgments. The latter criterion is crucial to establish a relationship between value judgments and factual arguments. Non-basic judgments may be revised based on a factual argument, while the basic ones may not. "The distinction is a simple one and lies at the root of relevance of factual consideration in ethical debates" (Sen 1967, 51).[26]

This classification allows Sen to clarify the relationship between preferences and choices. He takes the statement 'I prefer X to Y' to explain the relationship between preferring and choosing. He highlights three important elements of this statement: (1) It is a factual statement, because it shows someone's attitudes to X and Y. (2) This statement can represent a non-compulsive judgment, that is to say, someone agrees about choosing X, given a choice between X and Y, if there is no other judgment in his value system that gives a reason for choosing Y. (3) This statement can represent a compulsive judgment, which means, if nothing more is said, it can be supposed that there is no other judgment that gives a reason for choosing Y. But even if there are no other reasons for any different evaluative judgment giving a reason for choosing Y, other relevant facts may exist and have to be considered. Through these three essential elements, Sen demonstrates that it is "not self-contradictory" to ask: "I prefer X to Y, but what shall I do? (Sen 1967, 60-1).

As a result, even if an economic agent has a well settled preference, his value system or some relevant facts could be crucial for making his final choice between two states (X and Y). Besides, it is important to mention that the three classification criteria proposed by Sen are interconnected. They have therefore to be studied together in order to understand the relationship between preferences and choices explained above.

1.2.2.3 Comparison between Mongin's framework and Sen's framework

The main difference between these two approaches refers to the fact that Mongin tries to examine the role of value judgments of economists in their professional work and by contrast,

[26] The classification proposed by Sen (1967) will be explained in detail in section 1.3 of this chapter.

Sen tries to analyze the role of value judgments sometimes of economists and sometimes of economic agents. This difference will be the starting point of our analysis since the objective of each author is not the same, and the tools to develop their framework are significantly different.

The tools used by Mongin are based on the study of economic evaluations, which may be judgments of facts or judgments of value. He focuses on specifying the elements that make up the judgments of value (sentence, statement, utterance or inscription, and judgment). This first step in his analysis is essential to making the difference later among evaluations, prescriptions and obligations. The next step in Mongin's groundwork is to make a difference between thin and thick predicates, which are crucial to define the nature of evaluative predicates.

The tools used by Sen are based on Hare's meta-ethical framework, which assumes that an evaluative term has two meanings: prescriptive and descriptive. "Evaluative" judgments cannot thus be separated from a prescriptive meaning. Hare's distinction is essential for Sen because it allows him to establish the first classification: purely prescriptive and evaluative judgments.[27] Later on, he proposes two other classifications: compulsive and non-compulsive judgments, and basic and non-basic judgments. Sen establishes the connections between them in order to clarify the nature of ethical propositions.

Once the instruments used to shape the framework are presented, it is essential to know how each author uses it to demonstrate the role of value judgments, either in the case of the economist or in the case of the economic agent.

In Mongin's case we look at the following sentence made by a hypothetical development economist and taken up by him: 'If people are sufficiently well fed, and have access to basic education and health care, they do not live in a state of poverty'. Mongin considers poverty as a thick predicate because it is something undesirable (evaluative content), and at the same time it is a state of affairs that can be described by the concept of 'poverty line'. A diagnosis of the sub-statement 'they do not live in a state of poverty' thus leads Mongin to conclude that it is logically both evaluative and factual. It is assumed that the economist sincerely asserts the complete statement. The economist thus makes both a judgment of value and judgment of fact.

This example shows how Mongin uses his two-stage analysis to demonstrate how an economic statement can be qualified as positive and normative at the same time, which corroborates that

[27] Putnam criticizes Sen's classification because "it lies in squarely within the framework of Hare prescriptivism" (Putnam 2002, 70). Putnam affirms that Hare's belief that value terms are inherently motivational is unreasonable.

the authoritative criterion is not exclusive. Besides, we observe that the nature of predicates is a crucial element to determine if an economic statement is logically evaluative or logically factual.

In Sen's case, we examine two examples that show how his classification of value judgments is not precise concerning the question who makes the judgments of value: the economist or the economic agent. This lack of specificity of Sen shows that his goal is to study the nature of value judgments in order to clarify subjects such as: (a) "the relevance of factual considerations in ethical debates; (b) the problem of simultaneously holding some general moral principles that can conflict; and (c) the relationship between preferring and choosing". (Sen 1967, 62).

The first example refers to his first classification criterion, purely prescriptive and evaluative judgments. When someone says, "Your action last Sunday was courageous", not only the individual who makes this sentence commends the action, but he also describes it in a certain way. If instead, he says, "Your action was right", he almost commends it, but he may or may not mean to imply that it fitted in well with the accepted standards of the "right" behavior. So Sen concludes that the same sentence may express a purely prescriptive judgment and an evaluative judgment.

Based on this example, we can say three essential things: in the first place, the inseparability of evaluative and purely prescriptive judgments that can sometimes be presented in the same sentence. In the second place, the use of an ethical predicate as *courageous* may lead us to think that this sentence does not seem to be associated with the application of an economic policy. In the third place, the individual that makes the sentence might not be an economist, but an economic agent, which would not be far-fetched considering our last observation about the use of an ethical predicate.

The second example refers to Sen's third classification criterion, basic and non-basic judgments. Consider a person who makes the following value judgment, 'the government should not raise the money supply more than in proportion to the national output'. This would be based on a factual theory that relates money supply and output to inflation. However, this factual theory may be an object of debate and the person may move on to another judgment independent of the first factual assumption. So according to Sen, this person could continue to move on to more fundamental value judgments due to new factual assumptions, until he reaches a basic value judgment.

In contrast with the first example, we could assume that the above sentence seems associated with an economic policy recommendation, even if Sen did not identify at any time "the person" who makes the value judgment. This example shows us that economists have to deal with basic and non-basic value judgments, but more specifically that it is possible to discuss the validity of value judgments if economists focus on the underlying factual assumptions. In short, the two examples that we took from Sen's framework allow us to corroborate our thesis that the role of the economist and of the economic agent is not explicitly defined in his classification criteria.

On balance, it is evident that Mongin's framework is useful to analyze the role of value judgments of economists in their professional work. On the contrary, Sen's classification of value judgments should be carefully analyzed in order to determine if it is the economist or the economic agent who makes a value judgment. This lack of precision by Sen can be exploited in order to study the importance of the distinction between value judgments of economists and economic agents in his classifications. It allows us not only to use his criteria to analyze how an economic agent may justify his preferences, but also to have other tools different from Mongin's framework to study the role of value judgments of economists.

In the next section, we will analyze Sen's classification of value judgments in detail. We are going to assume that the criteria used by Sen can be used to examine the value judgments of economic agents. First, we are going to establish why the framework proposed by Sen is interesting in order to study the justification of the preferences of economic agents. Second, we present the definition of each criterion. Finally, we study the relationship between these criteria.

1.3 Value Judgments of Economic Agents in Welfare Economics

1.3.1 The Nature of Genuine Normative Reason of Economic Agents

> Within welfare economics, and to a large extent within political philosophy as well, the prevailing view has been that, at least for purposes of social justice, judgments of individual well-being should be based not on some set of "objective" value judgments taken as authoritative, or on the values of the person who is making the judgment, but rather on the preferences of the person whose welfare is being evaluated (Scanlon 1991, 23–4).

If, as Scanlon says in the quotation, the preferences of agents whose welfare is being evaluated are the basis of welfare economics, then it is important to know how these preferences are formed. It allows us to study the reasons that lead economic agents to choose between alternatives in a given situation. For example, a person may prefer traveling to Paris than to Kuala-Lumpur because he finds that Kuala Lumpur is not rich enough in cultural events. Someone else may prefer to go to Kuala Lumpur because he wants to know how the big modern Asian financial center looks like. This example allows us to see that different reasons underlie the preferences of economic agents.

The question to know how preferences are formed has been studied in depth by Dietrich and List (2013), who propose two conceptions of a reason: a motivating reason and a normative reason. The former is " a proposition that is motivationally relevant for the agent's preferences: if true of an alternative, it may affect the agent's actual preference for the given alternative vis-à-vis others" and the latter is "a proposition that is normatively relevant for those preferences: if true of an alternative, it may affect the preference the agent ought to have for that alternative vis-à-vis others" (Dietrich and List 2013, 5-6). We focus on the study of normative reasons, because they *justify* what economic agents do. Instead, the motivating reasons explain the actions of economic agents. If explaining why an economic agent chooses Paris instead of Kuala Lumpur is not trivial, justifying this choice may be a complicated exercise.[28]

In essence, these authors not only focus on "capturing the relationship between an agent's motivating reasons and his or her actual preferences" but also try to study the "relationship between an agent's normative reasons and the preferences he or she ought to have" (Dietrich and List 2013, 18). So they then take the same formal framework to analyze the reason-based explanations and the reason-based justifications of the agent's preferences. However, as it is well known, a motivating reason explaining a preference is not necessarily a normative reason justifying it. For this reason, we focus on the study of the justification of preferences through the classification of value judgments made by Sen.

The choice of Sen's framework to study the justification of preferences is based on the possibility that the classification of prescriptive judgments proposed by Sen (1967) may be enlightening in order to study the nature of normative reasons. This framework allows us to expand the normative perspective studied by Dietrich and List, who admits that "this formal

[28] See Maria Alvarez, *Reasons for Action: Justification, Motivation, Explanation* (The Stanford Encyclopedia of Philosophy, 2017).

analysis does not settle the question of what the right normative reasons are in any given context […]" (Dietrich and List 2013, 18).[29]

The analysis of these two authors is useful to determine how we can identify a motivating reason and a normative reason. They propose three clear features that determine the way a reason becomes motivating, but their argumentation is not clear enough when they tackle the determinants of a normative reason.[30]

We briefly present the three features that determine a motivating reason, which is part of the development followed by Dietrich and List to explain the preferences and choices of economic agents within the framework of a positive theory of rational choice. A motivating reason reaches motivational relevance for the agent's preferences when: 1) "in his or her conceptualization of the world in the relevant context, the agent distinguishes between those alternatives of which the proposition is true and those of which it isn't" (Dietrich and List 2013, 7). However, the abstract understanding of a proposition is necessary, but not sufficient to gain motivational relevance that affects the agent's preferences. For this reason, a motivating reason attains motivational relevance when: 2) "the agent qualitatively understands it" (Dietrich and List 2013, 8). Once the agent has understood abstractly and qualitatively some propositions, only some of them are 'typically salient for the agent'. Thus, 3) "the idea is that a proposition becomes motivating for an agent if he or she focuses on it actively or uses it as preferences-formation heuristic or criterion" (Dietrich and List 2013, 8).

These motivating reasons may be affected by normative reflections in case the economic agent is in a conscious, deliberate process of reviewing his or her preferences given a set of alternatives. The authors suggest that the third feature described above is the only one that best captures the role of normative reflection. The following quotation clarifies this point:

> An agent can deliberately interrogate him- or herself about which propositions are genuine normative reasons for him or her in relation to some alternatives, and thereby exercise some influence over which reasons come to motivate him or her (Dietrich and List 2013, 21).

The adjective 'genuine' to describe a normative reason is enlightened by the origin of this word. According to the Oxford Dictionary it comes from the Latin *genuinus*, which refers to the

[29] The formal analysis refers to "a formal calculus for linking normative reasons with ideal preferences, based on some underlying weighing relation" (Dietrich and List 2013, 19).
[30] Their analysis is justified when they define their primary aim to "develop a positive theory of rational choice" (Dietrich and List 2013, 7).

Roman custom of a father acknowledging the paternity of a newborn child by placing him or her on his knee. The motivating reasons of an agent can be influenced by an authentic (from the Greek *authentikos*) reflection that is inspired by the value system of the agent.

We are now going to study the nature of the value judgments of economic agents based on Sen's classification criteria. The idea is to analyze if these classification criteria allow us to shed light on the justification of preferences of economic agents.

1.3.2 Sen's Classification of Value Judgments

As we have already mentioned in sub-section 1.2.2.2 , Sen (1967) proposes three different criteria of classifying value judgments: (1) purely prescriptive and evaluative judgments; (2) compulsive and non-compulsive judgments; and (3) basic and non-basic judgments. We explain all three criteria in order to study the nature of value judgments and the relationship that may exist between them.

1.3.2.1 Purely Prescriptive & Evaluative Judgments

In the first criterion, a purely prescriptive judgment "intends to convey only an agreement to the underlying imperative, and not any factual information other than that necessary to express the imperative" and an evaluative judgment "implies not only the agreement to the underlying imperative, but also has a descriptive content" (Sen 1967, 46-7). For instance, if a person says: 'Female genital mutilation should be eradicated' and it means that he or she agrees with the imperative, 'Let us eradicate female genital mutilation', this can be taken as a purely prescriptive judgment. Conversely, if a person says, 'Female genital mutilation is cruel', the person not only agrees with the imperative, but also associates herself with the notion of cruelty of the procedures of female genital mutilation.

1.3.2.2 Compulsive & Non-Compulsive Judgments

In the second criterion, a compulsive judgment "implies a categorical imperative which can be held without reference to the rest of one's value system" and a non-compulsive judgment

"implies a hypothetical imperative which is conditional on there being no judgment in one's value system that gives a reason for a choice opposite to the non-compulsive judgment in question" (Sen 1967, 50). A compulsive judgment is presented in the statement 'It is better for you to go to the hospital now rather than to stay at home with this unbearable stomach-ache'. By contrast, a non-compulsive judgment is given by the example: 'Let me choose 'go to the opera' if I have no other reason to judge 'go to the movies' to be better'.

1.3.2.3 Basic and Non-Basic Judgments

In the third criterion, a basic judgment of a person means that there is "no conceivable revision of factual assumptions that can make him or her revise the judgment" and a non-basic judgment is conceived "if such revisions take place in the individual's value system" (Sen 1967, 50). To illustrate this distinction, we look at a person who says: 'A drop in household consumption measured at base year prices indicates a worse economic situation'. If he revises his judgment when someone proposes to him to consider certain factual circumstances, then this judgment can be considered as 'non-basic' in his value system. If, conversely, there is no situation where he regards adultery to be justifiable, then to not commit adultery is a basic value judgment in his system.

With regards to this distinction, it is important to consider three possible misunderstandings to be avoided, which Sen (1967) proposes. First, a revision of judgment is not the same as a revision of the wording of the statement expressing it. For instance, consider a conversation between Jean (J) and Boris (B). J: 'Mary should not eat like a vegetarian.' B: 'Mary is not a vegetarian. She is following the vegan trend.' J: 'OK, I say that Mary should not eat like a vegan.' The revision of the wording of the statement in the dialogue does not show Jean's judgment to be non-basic.

Second, factual revisions that can be allowable are not necessarily probable ones. We see it in the following dialogue. J: 'People should be allowed to eat as they like.' B: 'Even if it turned out that certain types of diets cause malnourishment?' J: 'Not in those cases, of course. But I don't think those situations are widespread.' Although they are unlikely to be malnourished due to a diet, the judgment is considered non-basic.

Third, "a value judgment may be conditional on certain factual circumstances" (Sen 1967, 51). For instance, consider the following judgment, 'On summer days, I should put on sunscreen'.

This is not shown to be non-basic by acknowledging that I will recommend something different for winter days, and by showing that I can make another recommendation even for summer days if, say, a sunscreen costs half my monthly income.

These three warnings of Sen allow us to understand the nature of his distinction, but it shows the difficulty of demonstrating when a value judgment can be called basic. In essence, "if the judgment expressed happened to be a 'basic' one in the value system of the person expressing it, then and only then could it be claimed that there can be no factual method of disputing the judgment" (Sen 1967, 52). For this reason, when Sen proposes this classification, he says that it could be plausible to demonstrate that some value judgments are non-basic, but it is not conceivable for those considered as basic.[31]

This criterion has a philosophical basis that is important to highlight. We refer more precisely to the distinction between intrinsic and extrinsic or instrumental value. The "intrinsic value of something is the value that that thing has for its own sake, while the instrumental value of something is for the sake of something else to which it is related in some way" (Zimmerman and Bradley 2019).

In this way, we assume that the concept of intrinsic value is related to a 'basic' judgment, because there is no way to find a factual method that changes the basic judgment of a person. That is to say, it is a value judgment that can be considered as good or bad for its own sake. Once the relationship between intrinsic value and 'basic' judgment is explained, we clarify why we assume that 'non-basic' judgments are related to the concept of extrinsic value. An extrinsic or instrumental value is derived from something else that is related to it, and a 'non-basic' judgment is precisely affected by factual assumptions that provoke a revision of this judgment. The difficulty is thus not in establishing if a value judgment is 'non-basic', but in establishing the fact that determines its basicness.

[31] Blaug (1998) argues that normative economics invokes two different types of value judgments, namely, the basic or non-basic value judgments proposed by Amartya Sen. As the most part of value judgments expressed in social questions are non-basic or impure, "a debate on value judgments can take the form of an appeal to facts, and that is all to the good because our standard methods for settling disputes about facts are less divisive than those for settling disputes about values" (Blaug 1992, 116).

1.3.3 Relationship between Sen's classification criteria

1.3.3.1 First criterion & Second criterion

The relationship between these two classification criteria implies that on the one hand, non-compulsive judgments are often 'evaluative', but it is not always the case. On the other hand, compulsive judgments are often 'purely prescriptive', but they are not always so. The table below shows the diverse associations that can be established between the two criteria.

Table 1. Examples of the relation between the first and the second criteria

Value judgment	Association
X should be chosen, not Y	Compulsive and Purely Prescriptive
X should be chosen, not lazy Y	Compulsive and Evaluative
X is, at least, more industrious than Y	Non-compulsive and Evaluative
X is, in some ways, more commendable than Y	Non-compulsive and Purely Prescriptive

(Sen, 1967, 62)

1.3.3.2 First criterion & Third criterion

According to Sen evaluative judgments must be non-basic, "because the prescriptive meanings do not exhaust their entire meaning, and there is some descriptive meaning also" (Sen 1967, 54). If we think of the female genital mutilation example, which was proposed in the definition, the person can conceive a factual revision about the cruelty of female genital mutilation and the judgment would change. With regards to prescriptive judgments, it cannot be said that they must be basic, because "even a purely prescriptive judgment can be based on some factual assumption relating it to some more fundamental value" (Sen 1967, 54). For instance, 'Female genital mutilation should be eradicated' can be based on a factual assumption relating to the case of at least 200 million women who are victims of this harmful practice in 31 countries

around the world.[32] The person may however move on to a more fundamental value judgment that justifies the eradication of female genital mutilation, such as considering that this form of violence and discrimination against women is a human rights violation.

1.3.3.3 Second criterion & Third criterion

The relationship between the compulsive/non-compulsive distinction and the basic/non-basic distinction shows us that in the individual value system not all the judgments can be basic and compulsive at the same time. Take a person who 'believes in' honesty. When he has only a choice between a situation where lying has serious consequences and a situation where lying has less serious consequences, he will choose the latter. He thus makes a non-compulsive judgment. Through this example we can see that if honesty is considered as a basic value judgment, then the choice of the individual might not be considered as a compulsive judgment. Nevertheless, if the person believes that lies can never allow being faithful, then he/she can put forward the principle of honesty as a compulsive judgment. But in this form, it will be non-basic because there is an underlying factual assumption given by the belief that 'lies can never allow being faithful'. In short, if a person believes in the principles of both honesty and faithfulness, then one of the two judgments must be either non-basic or non-compulsive (Sen 1967, 55-6).

To sum up, the three criteria proposed by Sen offer elements of an answer to the way how economic agents justify their preferences. We do not assume that one of these criteria is better than another in order to understand the role of value judgments in the justification of preferences, because as we have seen there is a relationship between the three classifications. Hence, economic agents can *justify* their preferences if there is a set of evaluative judgments that can play a role in their choice; if there are relevant facts that have to be considered; or if there are categorical or hypothetical imperatives to which the agents agree. Furthermore, the study of this framework allows us to understand the importance of the value system of economic agents, which will determine the nature of their judgments when facing a choice.

[32] Statistics can be found on the Unicef (2022) website.

1.4 Economists' Role: Qua economists and Qua citizens

When we study the role of values in the decisions that economists make, we could assume that economists are like the Greek god Janus who has two faces and represents among other things duality. In the case of our research this duality is in the relationship between economists in their professional work and economists as citizens. The importance of this double role lies in the economists' ability to separate the values that they have when they have one role or the other. We thus identify economists *qua economists* and economists *qua citizens* as Mongin (2006b) did in his article on "Value Judgments and Value Neutrality in Economics".

In essence, if economists cannot separate their values according to the role that they assume, then the principle of axiological neutrality proposed by Weber (1919 [1965]) is broken because economists *qua economists* do not repress the value judgments that arise spontaneously in their role as economists *qua citizens*.

By contrast, authors like Gunnar Myrdal consider that the economist *qua economist* and the economist as *qua citizen* have the same value judgments. "The interests can never be purely scientific. They are choices, the products of our valuations. Without valuations we have no interest, no sense of relevance or of significant and, consequently, no object" (Myrdal 1958, 51).

We observe that there are two opposing ideas about the way how economists face value judgments in decision-making. We first take Sen's framework to analyze the economists' duality. Then we study Mongin's (2006b) theses about the role of value judgments in economics which is useful to show that there is an intermediate position between the points defended by Weber and Myrdal.

1.4.1 Sen's framework as a multipurpose tool

When we looked at basic and non-basic judgments as tools to *justify* the economic agent's preferences, we realized that there is nothing that prevents us from using the same framework to analyze the possibility that economists can make value judgments in their profession. Sen (1967) helps us to understand the entanglement between values and facts which is rejected by

Robbins (1935 [1984]) (see sub-section 1.5.1). Sen affirms that the contrast between ethics and economics "would hold if ethics dealt only with basic judgments". But, "the assumption that judgments on certain specific fields must be basic in everyone's value system, does not seem to be particularly realistic" (Sen 1967, 53).

Sen thus proposes two possibilities to test the basicness of a value judgment. The first test is to ask the person concerned if his or her judgment is basic. "But since no one would have had occasion to consider all conceivable alternative factual circumstances and to decide whether in any of the cases he would change the judgment or not, his answer to the question may not be conclusive" (Sen 1967, 53). The second test is to ask the person to consider a series of appropriate revisions of factual assumptions, but "this process never establishes basicness, though it can establish that the judgment is not non-basic in any obviously relevant way" (Sen 1967, 53).

These tests of basicness can be useful to demonstrate that some value judgments are non-basic, but they cannot provide a satisfactory answer to demonstrate that a value judgment is basic. In the case where a value judgment is subject to a disagreement between two individuals, it is possible to have "a scientific discussion on the validity of the value judgment by examining the scientific truth of the underlying factual premise" (Sen 1967, 52). The importance of this possibility is that value judgments cannot rule out a debate among economists. This is a criticism of Robbins' framework and an interesting clue to analyze the role of values of the economist *qua economist*.

1.4.2 Mongin's framework: The Value Neutrality Problem

The typology of the neutrality theses presented by Mongin (2006b) is a useful framework to analyze the different debates in the twentieth century about the neutrality of economists. He proposes four theses (Table 2) that are characterized by the degree of neutrality that an economist may have at the moment of carrying out his or her professional work.

Table 2. Four theses to study the value neutrality problem

Degree of neutrality	Theses
Strong Neutrality (SN)	Economists should always abstain from making value judgments.
Weak Neutrality (WN)	Economists might, and even should, make value judgments.
Strong Non- Neutrality (SNN)	Economists may not abstain from making value judgments.
Weak Non-Neutrality (WNN)	Economists might, might and should, or might not, make value judgments, depending on the case at hand.

It is important that Mongin defends the WNN thesis based on his two-stage analysis (see sub-section 1.2.2.1). He summarizes the process:

> Technically, the weak non-neutrality thesis 4 depends on splitting the traditional concept of a value judgment into a statement and a judgment proper. This leads to a two-stage procedure to discriminate between judgments of value and other judgments. The first stage scrutinizes the statement and raises only one question, but a difficult one: what kind of predicates does the statement contain? […] The second stage, which is of particular interest for statements containing evaluative predicates, involves again only one question, though again a difficult one: is the economist responsible for the judgment associated with the statement? […] (Mongin 2006b, 261).

The role of predicates in the first stage is crucial in order to know if the statement is logically evaluative or logically factual. He establishes an important difference between thin predicates (like *good*) that are only evaluative, and thick predicates (like *relevant*) that are evaluative and descriptive at the same time. Value judgments are not easy to identify and difficult to separate from judgments of fact. It is thus not possible to consider positive and normative economics as two different areas.

Mongin looks at the weaknesses of the WN thesis, because this thesis is between the two extreme theses, SN and SNN, and also because it has been put forward by economists within new welfare economics and social choice theory. He analyzes what he calls the 'containment' claim, which is the basis of WN thesis: "(i) the value judgments that economists have a right to make are few in number; (ii) the value judgments are easy to discover, and (iii) these judgments

can be logically and practically separated from the judgments of fact also made by economists" (Mongin 2006b, 259).

To demonstrate that the 'containment' claim cannot be guaranteed, Mongin examines three cases of economic statements, which deny the authoritative criterion. First, he takes Weber's analysis of the trade unionist to explain the indeterminacy between the positive and the normative. Second, he examines how thick evaluative predicates function semantically, which can be crucial to determine whether statements would qualify as being positive and normative at the same time. Finally, he takes an example in which the division is categorical but unintuitive.

Mongin takes Weber's trade unionist case because he considers that it is a 'classical case' that describes the weak neutrality thesis well. This case explains how the social scientist tries to explain the trade unionist's value commitments but without interfering with any of his values. Mongin's criticism refers to the fact that the 'trade unionist' is considered an *ideal-typical* individual, which comes into tension with axiological neutrality. This is illustrated in the following example:

> A trade unionist makes value judgments V1, . . ., Vn. Each of these judgments corresponds to a statement made by the representative individual, and a set of statements may or may not be logically consistent. If it is not, the question arises: which of the statements corresponding to V1, . . ., Vn should be taken out in order to restore consistency? Supposing that consistency has been restored (or was there from the start), another question arises: how is the set to be expanded in order to provide a sufficiently complete picture? Weber assumes that logic is all that the social scientist needs in order to adjust the initial V1, . . ., Vn. (Mongin 2006b, 277).

Mongin doubts that logic is the only element that the social scientist has to consider in order to adjust the initial value judgments made by the trade unionist. Social scientists' evaluations may be considered restoring the consistency of the value judgments, which may pervade their ideal-typical restorations of value positions. In brief, through Weber's classic case, Mongin demonstrates that there can be an indeterminacy between positive and normative, which is one of the reasons why he does not support the weak neutrality thesis.

If we study the other two cases that Mongin proposes to demonstrate that it is not possible to make a clear division between the normative and positive aspects in economics, we will find the analysis of the thick predicate 'poverty', which has already been examined in the sub-

section 1.2.2.3; and also the case where 'the authoritative criterion sometimes involve a counterintuitive division between the positive and the normative'.

If we suppose that an economist who investigates a value judgment made by an economic agent, not only supports this value judgment, but also makes it clear that he does so; then according to the authoritative criterion a normative statement was made. Mongin claims that reaching this conclusion can be "completely unnatural." He thus proposes the following example to show it: 'The Muslim part of the population thinks that it is despicable to eat pork, which affects the pork consumption in East London.' At first, if we follow the two-stage analysis and the authoritative criterion, explained in 1.2.2.1, this sentence is part of positive economics. But if the economist adds: 'I disapprove of eating pork.' This second sentence brings more information. Then, the sub-statement 'It is despicable to eat pork' is logically evaluative and it is possible to say that the economist sincerely accepts it. So the two-stage analysis leads to say that the economist has made a value judgment; and, according to the authoritative criterion, we can finally say that the first sentence is normative (Mongin 2006b, 279).

Mongin however supposes that the economist expresses the second sentence as an incidental warning. This could be because the economist is a Muslim himself, and he is worried that his religion may affect his interpretation of facts. So this second sentence is given to precisely preserve the positive character of the first sentence. In essence, we observe that it is not because the economist makes a value judgment that his work can be considered as normative.[33] To sum up, Mongin points out the limits of the weak neutrality thesis, and the impossibility to make a clear distinction between normative and positive economics.

For the theses with extreme positions, Mongin uses the first stage of his two-stage analysis to reject them. On the one hand, the SNN thesis that is defended by Myrdal can be summarized as: "one cannot be seriously interested in an object, or make a selection relative to that object, without evaluating some or all of its features" (Mongin 2006b, 274). In contrast with the last proposition, Mongin says that is not possible that everything in economics is evaluative, because not all the predicates in economics are evaluative. We observe the importance of predicates to determinate if a sentence is logically evaluative, logically factual or logically neutral.

[33] This example about the pork consumption in East London brings to light the importance of semantics for the positive-normative distinction.

On the other hand, the SN thesis that is supported by Robbins is based on the dissociation between economics and ethics, which will be explained in sub-section 1.5.1. Mongin claims that Robbins and the defenders of this thesis use the Weberian formulation to define the problem of value neutrality, but its solution is presented in Humean terms. Mongin's criticism lies in the idea that Hume's thesis (one cannot infer an "ought to" from an "is") could be set aside from positive economics only in the prescriptive part of normative economics and not in the evaluative part. Besides, Mongin claims that making an evaluation is not making an ethical evaluation, because not all the evaluative predicates of economics are ethical. This proposition avoids Robbins' dichotomy between economics and ethics.[34]

To summarize, according to Mongin's two-stages analysis the weak non-neutrality thesis is the most appropriate to study the role of values of economists in their professional work. This implies on the one hand giving importance to the nature of predicates in the sentences made by economists, which will be crucial to determine if the sentence is logically evaluative or logically factual. On the other hand, it is difficult to make a clear distinction between normative and positive economics. For this reason, Mongin not only discards the theses that are at the extremes (SN and SNN) but also the WN thesis, which takes normative and positive economics as two distinct branches of economics.

Once we have established the theoretical framework to analyze the distinction between the value judgments of economists and those of economic agents, we are going to study the importance of this distinction. We take as a case study Harsanyi's approach to ICU's. The aim is to show that it is not possible to 'objectify' subjective attitudes of economic agents, and that economists cannot avoid value judgments when they recommend an action for the benefit of the society.

1.5 Interpersonal Comparisons of Utility as a Tool: the Role of Values of Economists and Economic Agents in Economics

Interpersonal comparisons of utility (ICU's) have been at the center of the debate in welfare economics. The debate is focused on the possibility or impossibility of making interpersonal

[34] Mongin proposes as an example the predicate: rational. "Rational is clearly evaluative in economics but does not belong to the same group as *just* because it does not have any direct ethical counterpart. Accordingly, it can at most be invested with an indirect ethical relevance […] The meaning of this predicate depends on restrictions, some of them stipulated, others implicit, which load it with a descriptive content" (Mongin 2006b, 271).

comparisons of utility in welfare economics. We have two opposite fields, the first welfare economics which allows ICU's, and the 'new welfare economics' which forbids them.

In the first welfare economics, ICU's can play an essential role in determining economic policies to improve people's well-being. The way ICU's are made reveals a normative conception. If comparisons of utility are made by sum-total calculation, the calculation is utilitarian. If the comparisons are made by considering the order of utilities and considering only the lowest utility, the calculation is egalitarian. To illustrate the comparisons of utility, Baujard (2017) takes the consumer surplus definition proposed by Alfred Marshall, which is important to evaluate the well-being of an individual and a society. The consumer surplus is based on the cardinal evaluation of individual utilities. ICU's are possible because economic problems can affect all classes of society in a similar way, which leads one to avoid the problem of heterogenous agents.[35]

There are two important criticisms of the first welfare economics, which are well identified by Mongin (1999). The first criticism is based on Robbins (1935 [1984]), who rejects the idea of making ICU's because these comparisons are based on value judgments and not on empirical facts. The second criticism is based on Paretian theory, which attacks ICU's because, the economist cannot match the cardinal evaluations of individual utilities (comparisons at intrapersonal level) with possible experience.[36] The two criticisms led to the emergence of the 'new welfare economics', which is characterized by the use of ordinal utility functions. These functions allow one to represent the choice of individuals, but not directly their well-being. Besides, in the 'new welfare economics', as ICU's are rejected, the only way to compare social states across a population is through the Pareto criterion.

Nevertheless, value judgments cannot be avoided even when using the Pareto criterion, because when economists consider that an economic situation x is Pareto-wise better than an economic situation y, it entails that x not only provides greater well-being but is also better. Mongin (1999) supports this reasoning by an analysis of the economist's reasoning within the framework of 'new welfare economics':

> The more an individual's preferences are satisfied, the greater his economic well-being; the greater the economic well-being of an individual, the greater his well-being considered as a whole; the greater the well-being of all individuals taken

[35] See Elster and Roemer (1991) for studying the relevance of interpersonal comparisons of utility in welfare economics.

[36] Pareto rejects interpersonal comparisons of utility because he thinks that these comparisons belong to a "pseudoscience" (Pareto, 1909, comment by Mongin 1999).

> together, the greater the well-being of the society to which they belong (this is the "individualistic" stage of reasoning); and ultimately, the better is this society (Mongin 1999, 540) (My translation)

We observe how the fact that there is an increase in the well-being of an individual contributes to a better society reflects a value judgment, because we would be in a more just society. So, even if the 'new welfare economics' avoid value judgments involved in making interpersonal comparisons of utility, it cannot entirely prevent value judgments.

The debate among the supporters of the first welfare economics and the 'new welfare economics' is focused on the possibility of making value judgments in the study of the well-being of individuals. On the one hand, using ICU's allows legitimizing the economic policies proposed by economists. On the other hand, ruling out ICU's allows legitimizing the scientific character of economics. Although these two views seem irreconcilable, there is an author who tries to show that it is possible to make ICU's, without resorting to value judgments.

We are referring to Harsanyi, who justifies the use of ICU's through an additive cardinal social welfare function derived from the von Neumman-Morgenstern axioms. From these axioms, Harsanyi assumes rational behavior, which is essential for determining the rational decision of an impartial observer choosing lotteries describing the situation of individuals.[37]

The analysis of Harsanyi's approach to ICU's allows us to study the importance of making a distinction between the role of values of economists and that of economic agents in economics. This distinction has been hidden behind the debate about making ICU's or not in welfare economics. The objective of this section is thus to show how this distinction is necessary and should be part of the debate.

We first study Robbins' reluctance to ICU's. He is one of the big supporters of the 'new welfare economics.' The difference he proposes between economic science and economic policy is crucial in order to understand the way to legitimize the scientific character of economics. We then present Harsanyi's impartial observer theorem to explain one approach to making ICU's objectively. We study three criticisms of Harsanyi's proposition: (1) Broome's criticism of the causal variables proposition; (2) Mongin's criticism of the uniformity of extended preferences;

[37] It is important to note that there are different forms of ICU's. We have interpersonal comparisons of utility levels, based on ordinal comparisons of real numbers, and interpersonal comparisons of utility differences, based on cardinal comparisons of utility functions. See Hammond (1977; 1991) for a comparison of these two forms of ICU's.

and (3) Mackay's criticism of the problem of the experimental contamination of the impartial observer.

1.5.1 Robbins' reluctance to Interpersonal Comparisons of Utility

Robbins (1935 [1984]) in the last chapter of *An Essay on the Nature and Significance of Economic Science* establishes the importance of the existence of 'some analytical apparatus' in economics, which allows us to understand the incompatibilities that are present among economic agents and to examine the long-term consequences of different policies. Robbins thinks that economic analysis is an essential tool to judge an economic phenomenon rationally, but that it is not useful to choose the final economic policy:

> Faced with the problem of deciding between this and that, we are not entitled to look to Economics for the ultimate decision. There is nothing in Economics which relieves us the obligation to choose. There is nothing in any kind of science which can decide the ultimate problem of preference […] For rationality in choice is nothing more and nothing less than choice with complete awareness of the alternatives rejected. And it is just here that Economics acquires its practical significance (Robbins 1935 [1984], 152).

This practical significance of economics means that economists are analytical technicians that put into practice their knowledge to study economic problems rationally without entangling their value judgments. This conception highlights the status of economics as a science, and it implies that economists have to provide an analysis that can be useful for policy makers. But that does not mean to say that economists may express their preferences about alternative policies, because values are not scientifically verifiable.

A quick read of Robbins could lead us to think that economists should not consider ethical questions, but this is not clearly the idea defended by Robbins. On the contrary, what he proposes is that there is no logical connection between ascertainable facts and value judgments, a position which is based on the dichotomy proposed by Hume (1739 [1960]), who assumes that there are two types of judgments: 'is' judgments and 'ought' judgments. The importance of this distinction lies in the well-known law of Hume, which states that it is not possible to infer an 'ought' judgment from an 'is' judgment.

One of the repercussions of the propositions made by Robbins is the difference between economic science and economic policy, because it allows one to establish that the role of

economists is to use their knowledge to analyze and interpret economic phenomena in order to propose different policy alternatives. It is thus the policy makers who will be in charge of deciding if the ideas proposed by economists are good or bad for society. It is evident that this position puts the role of economists as advisors into question. Hence it is important to have in mind in what context economists are: that of economic science or of economic policy:

> Robbins's claim was actually different, namely that economists cannot give policy advice qua practitioners of economic science. Nothing prevented them from giving policy advice qua practitioners of political economy, another facet of their professional capacity (Scarantino 2009, 460).

Another crucial repercussion of Robbins's essay is related to interpersonal comparisons of utility, which he regards as subjective and inherently normative. This conception is logical with regards to the difference between economic science and economic policy proposed by Robbins because ICU's allow one to evaluate the pros and cons that policies might cause to society:

> The theory of exchange […] does not assume that, at any point, it is necessary to compare the satisfaction which I get from the spending of 6d. on bread with the satisfaction which the Baker gets by receiving it […] It is a comparison which is never needed in the theory of equilibrium and which is never implied by the assumptions of that theory. It is a comparison which necessarily falls outside the scope of any positive science (Robbins 1935 [1984], 138-9).

In contrast to Harsanyi, who justifies the interpersonal comparisons of utility through his proposal of causal variables, Robbins rejects ICU's because they belong to the field of normative economics. As a result, we may not make ICU's without making value judgments and this is crucial in order to define the role of economists as analytical technicians. Robbins does not deny that "economists should deliver themselves on ethical questions", but he regards that "the validity of assumptions relating to the value of what exists or what may exist is not a matter of scientific verification" (Robbins 1935 [1984], 149-50)

This reluctance to use ICU's in economic science had a great impact on welfare economics and social choice theory. In the case of welfare economics, it meant an attack on its basis and led to the emergence of "new welfare economics", which is limited to identifying Pareto efficient outcomes. In the case of social choice theory the rejection of ICU's is crucial in the derivation of Arrow's (1951) impossibility result, because the informational basis is restrained. Sen (1970 [2017]), inspired by the goal of overcoming Arrow's impossibility theorem, proposes the inclusion of interpersonal comparisons of utility in the informational basis of the social choice

theory, which implies that value judgments are unavoidable in the Collective Choice Rule (CCR).

1.5.2 Harsanyi's Impartial Observer Theorem

Before formally presenting the theorem, it is important to understand the concept of extended preferences proposed by Harsanyi (1955, 1977), because as Mongin says: "extended preferences is a natural tool to model the impartial observer" (Mongin 2001a, 11). Harsanyi (1977) gives an example to introduce this concept:

> Thus let A be a social situation where all individuals' diets consist mainly of fish and let B be a social situation where all individuals' diets consist mainly of meat. Suppose that individual i has a mild preference for fish, while individual j has a very strong preference for meat (with a violent distaste for fish). Then individual i, his taste Pi being what it is, will obviously prefer fish to meat, which means that he will prefer $[Ai, Pi]$ to $[Bi, Pi]$. But he will presumably also recognize that it is better (less inconvenient) to eat meat with a mild distaste for meat than it is to eat fish with a strong distaste for fish. Therefore, he will prefer $[Bi, Pi]$ to $[Aj, Pj]$. In terms of the language of interpersonal utility comparison, he will recognize that j would derive more disutility (i.e., would derive less utility) from eating fish than (i) himself would derive from eating meat. (Harsanyi 1977, 52-3)

In this example, we identify some elements that are essential in order to understand the extended preferences notion: (i) the social situation that represents the context that surrounds the individuals, (ii) the preferences between ideal alternatives, (iii) the empathy that one individual has for the other.

The first element (i) of the individuals' dietary regime is influenced by causal variables, namely the preference i has for fish ($[Ai, Pi]$ to $[Bi, Pi]$), and the possibility that i could have other preferences while still being i ($[Bi, Pi]$ to $[Aj, Pj]$). Harsanyi proposes that these causal variables are considered as objectives and they are needed to explain "the subjective attitudes of individuals (including their preferences)" (Harsanyi 1977, 58).

The second element (ii) of ideal alternatives is important in order to establish a difference between an individual's personal and extended preferences. According to Harsanyi, "an individual's personal preferences will manifest themselves both in this actual choice behavior and his verbal statements of preferences, whereas in general his extended

preferences can manifest themselves only in his verbal statements" (Harsanyi 1977, 53). The ideal alternatives are thus represented by verbal statements, but not by preferences between real options.

The third element (iii) of empathy is at the base of the extended preferences. Sometimes it is confused with sympathy. But these terms are quite different. According to the Oxford Dictionary, empathy is "the ability to understand and share the feelings of another" (individual i recognizes that j would derive less utility from eating fish than i himself would derive from eating meat). By contrast, sympathy means the "feeling of pity and sorrow for someone else's misfortune".

More specifically, Harsanyi explains the term 'imaginative empathy' as a process necessary to make interpersonal comparisons of utility and puts it at the base of his definition of extended preferences. This process allows him to create an objective framework of interpersonal comparisons of utility, which must be based on the causal variables that explain subjective attitudes and personal preferences. These causal variables, according to Harsanyi (1955, 1977), are factual information such as the individuals' psychological, biological, social and cultural characteristics and the general psychological laws governing human behavior.

We cannot however discard the sympathy characteristic that is mentioned by Harsanyi (1977) in different passages of his paper when he talks about the impartial observer. This observer has not only the capacity to have an impartial moral point of view at the moment of judging the welfare of two individuals, but also the ability to correlate some of his judgments of preferences with those of the observed individual positively.

We now look at the consistency axioms for extended preferences proposed by Harsanyi (1977, 54–5):

- Axiom 1. The rationality of individual i's extended preferences. Extended preferences of individual i among mixed extended alternatives in set ψ satisfy the axioms for rational behavior under risk.[38]

[38] Harsanyi considers risky prospects or probability mixtures "whose components are pure extended alternatives, possibly belonging to two or more different individuals, for example, probability mixtures of $[A_j, P_j]$ and $[B_k, P_k]$"(Harsanyi 1977, 53-). These are called mixed extended alternatives, and the set of these latter is called ψ.

This axiom is required to establish that, in making moral value judgments, individual i will attempt to maximize the expected utility $W_i(A)$ (Harsanyi 1977, 54).

$$W_i(A) = \frac{1}{n}\sum_{j=1}^{n} U_j(A) \qquad (1)$$

Equation (1) means that individual i has an equal probability $1/n$ of swapping places with any other individual j $(j = 1, \dots, i, \dots, n)$ obtaining the utility $U_j(A)$, which represents individual j's utility level in a social situation A.

- Axiom 2. Agreement between individual i's extended preferences and any other individual's personal preferences. Let A_j and B_j be two risky prospects in set Y, and let j be a given individual $(j = 1, \dots, i \dots, n)$. Let's suppose that $A_j \gtrsim B_j$ in terms of the personal preferences of individual j. Thus also $[A_j, P_j] \gtrsim [B_j, P_j]$ in terms of extended preferences of individual i.
-

This axiom has two cases:

- The case where $j = i$
  "To say that individual i prefers $[A_i,\ P_i]$ to $[B_i,\ P_i]$ is the same thing as saying that, his subjective attitudes P_i being what they are, he prefers position A_i to position B_i" (Harsanyi 1977, 54).
- The case where $j \neq i$
  "The axiom… is a restatement of the principle of acceptance (consumers' sovereignty). It expresses the fact that individual i will evaluate the personal position of another individual j in terms of j's own personal preferences" (Harsanyi 1977, 54).

Considering Axiom 1, an extended-utility function can be defined $V_i = V_i[A_j, P_j]$ "for individual i over all extended alternatives $[A_j, P_j]$ with $j = 1, \dots, i, \dots, n$ and over all probability mixtures of such alternatives" (Harsanyi 1977, 54).[39]

Considering Axiom 2, a utility function U_j can be defined for each individual j as follows:

[39] Harsanyi considers risky prospects or probability mixtures "whose components are different individual positions A_j, B_j, ... of the same individual j (Harsanyi 1977, 53).

$$U_j(A_j) = V_i[A_j, P_j] \text{ for all } A_j \in Y \quad (2)$$

Harsanyi assumes that V_i and U_j have the expected utility property and therefore the utility function of individual j is a von Neumann-Morgenstein utility.[40]

Harsanyi concludes his exposition of the consistency axioms for extended preferences by saying that:

> In making moral value judgments individual i will treat any social situation A as if it were an equiprobability mixture of the n extended alternatives $[A_1, P_1], \dots, [A_n, P_n]$ (Harsanyi 1977, 55).

From these two axioms we see two important principles for the formulation of Harsanyi's Impartial Observer theorem: the Equal Chance principle (EC) and the Consideration of Others principle (CO). The former connects moral value judgments to extended preferences and the latter extended to actual preferences.

Another principle to explain the Impartial Observer theorem is the Uniformity of Extended Preference (UEP). Harsanyi says this principle consists in assuming that "the ordinary utility functions of two individuals i and j may be quite different, but their extended utility functions will be identical" (Harsanyi 1977, 59). To understand this principle better, we propose an example inspired by Harsanyi: two friends, Bastien (who likes high-tech advice) and Patrice (who likes reading), compare their satisfaction level. Bastien compares the utility that he would derive from a new smartphone with the utility that Patrice would derive from a new Kindle (a tablet only for reading). At one point, Bastien must ask himself what utility he would derive from a Kindle if he had taken up reading as a regular hobby as Patrice has done, and if he had his friend's speed-reading skill.

Having identified these three principles we can present Harsanyi's Impartial Observer Theorem following Mongin's (2001a) method:
Assume that for each individual, the extended preferences relation satisfies the Von Neumann-Morgenstern (VNM) axioms, and that (EC), (CO) and (UEP) hold. Then there exists a common utility representation W of the individuals' moral preferences for social situation A, such that: for all $a \in A$, that is formally represented by equation (1).

[40] A given utility function U has the expected-utility property "if and only if $U(C) = U(A, p; B, 1-p) = p \cdot U(A) + (1-p) \cdot U(B)$ […] a utility function U having […] this property is called a von Neumann-Morgenstern utility function" (Harsanyi 1977, 32).
See John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 3rd ed. (London: Princeton University Press, (1953)).

1.5.3 The implications of Impartial Observer Theorem

We identify two important implications of this theorem: 1) Anybody can assume the role of an impartial observer; and 2) interpersonal comparisons of utility can be "objective" thanks to the identical extended preferences of the different observers.

The first implication may seem trivial when we study the theorem. Still, we think that if we identify who is the impartial observer, we may know who is 'objective' when making value judgments. Harsanyi's view concerning the anonymity of the observer is clear: "Indeed every person is expected to follow such impartial criteria when he makes moral value judgments" (Harsanyi 1977, 48). The impartial observer can thus be an expert who looks after social interests or any individual in society, depending on the context under study.

The importance of determining who is assumed to be an outsider observer is a question that makes sense when we aim to study the role that value judgments may play in economics. On the one hand is the case where economists try to recommend a specific action for the benefit of general social interests. On the other hand is the case of studying where value judgments may influence the social decisions of economic agents.

Consequently, if we want to study the differences between the value judgments made by economists and those made by economic agents, we cannot do it within the framework defined by Harsanyi. This distinction is not important for him, because he is focused on demonstrating that interpersonal comparisons of utility have a 'logical basis', which allows one to make these comparisons without the intervention of value judgments based on ethical or political postulates. His purposes are clear:

> Professor Robbins is clearly right when he maintains that propositions which purport to be interpersonal comparisons of utility often contain a purely conventional element based on ethical or political value judgments […] But in order to avoid confusion, such propositions based on ethical or political restrictive postulates must be clearly distinguished from interpersonal comparisons of utility without a conventional element of this kind (Harsanyi 1955, 320).

This conventional element refers to 'subjective' preferences, which express the actual preferences of individuals, "showing an egoistic attitude in the case of an egoist and an altruist attitude in the case of an altruist" (Harsanyi 1955, 315). By contrast, the 'ethical' preferences

express what individuals prefer or would prefer based on impersonal social considerations only. The existence of an impartial observer guarantees this detached attitude. Thus interpersonal comparisons of utility are possible because Harsanyi assumes that "individuals' choice behavior and preferences are governed by the same basic psychological laws" (Harsanyi 1977, 58).

Even if Harsanyi has diverted the impossibility of making interpersonal comparisons of utility, his methodology has been criticized by different economists who think that it is not possible to 'objectify' even value judgments. In the next sub-section we concentrate on the criticisms made by Broome (1993) and Mongin (2001a) of the methodology followed by Harsanyi. We also analyze Mackay's approach (1986) to the impartial observer's identity.

1.5.4 Criticisms of the Extended Preferences proposition: opening a Pandora's box

If something is evident in the last sub-section, it is that Harsanyi attempts to overcome the value judgments issue in interpersonal comparisons of utility. His main achievement is to keep the Pandora's box that represents value judgments in economics closed. We now study how value judgments cannot be put aside when we make interpersonal comparisons of utility. This is useful to understand why it has been so difficult to include subjective preferences in economics. Our aim is to concentrate on the importance of value judgments when we make interpersonal comparisons of utility and also on the actors that make these comparisons.

1.5.4.1 Broome's criticism

Broome's criticism goes to the heart of Harsanyi's causal proposition because for him it is not possible that everyone has the same extended preferences. Broome's objection is based precisely on the role of value judgments of individuals that Harsanyi had avoided with his hypothesis that the same psychological laws govern all individuals Broome clarifies his criticism as follows:

> My preference between the life of an academic and the life of a financial adviser differs from the financial adviser's preference. It is implausible that this is because of some irrationality or lack of information on the part of one of us. It is because we have different values. I value opportunities to add to knowledge; she values an opulent life-style. I very much doubt that either of these values is irrational, ill-informed or even objectively wrong (Broome 1993, 67).

We can see how Broome calls into question the preference-satisfaction condition, which says that one alternative would be better for an individual than another if this individual prefers it and if this preference is rational and well-informed. Thus for Broome, the difference between the two individuals' preferences lies in values, and this hypothesis erases the possibility that the same causal variables determine an individual's preferences.

The claim that Broome makes is that "a cause of preference must not be mistaken for an object of preference" (Broome 1993, 59). Formally, it is expressed as follows: $U(x;\phi)$ represents preferences for objects x (the amount of comfort and excitement in life at a particular time), which are determined by causes ϕ (age at the time). Another function $V(x;\phi)$ represents the same preferences and the same causal variable if and only if $V(x;\phi) = f(U(x;\phi),\phi)$, where f is increasing in its first argument. This transformation means that a change in ϕ that increases U need not increase V.

For instance, Patrice has preferences for two alternatives: family traveling at a particular time of his life (a) and adventure traveling (l), which are represented by the following function:

$$U(a,l) = (\theta)\log(a) + (1-\theta)\log(l) \qquad (3)$$

where θ is a parameter between 0 and 1. An increasing transformation may represent these same preferences:

$$V(a,l) = \log(a) + (1/\theta - 1)\log(l) \qquad (4)$$

We suppose that there is a causal variable that explains these preferences. Patrice's preferences for adventure traveling and family traveling depend on his age at the time. We assume that the parameter θ is represented by $t/100$, where t is his age. The utility function is then:

$$U(a,l;t) = (t/100)\log(a) + (1-t/100)\log(l) \qquad (5)$$

In this causal function, the role of the causal and object variables are quite different. Therefore, if a is greater than l, U increases with t. This does not mean however that Patrice likes getting older. His preferences about age are not represented in the function. If we substitute $t/100$ in equation (4), this point is reinforced according to Broome.

$$V(a,l;t) = \log(a) + (100/t - 1)\log(l) \qquad (6)$$

V is not an increasing transformation of U if the arguments of U are taken to be a, l and t. Thus, while U rises with t inasmuch as a is greater than l, this is not the case for V. V shrinks with t given that l is greater than 1.

This example shows us in a more intuitive way the hypothesis defended by Broome and it leads us to reconsider the proposition of coincidence of individuals' extended preferences, which is based on causal variables. It means that interpersonal comparisons of utility cannot be derived from extended preferences, and that the role of value judgments cannot be put aside as Harsanyi has done. For this reason, Broome recognizes that each individual has his own particular values, which are difficult to give up, even if an individual is forming his preferences for lives in which he would not have those values (Broome 1993, 65).[41]

1.5.4.2 Mongin's criticism

Mongin's criticism is focused on the questioning of the principle of the Uniformity of Extended Preference (UEP), which is necessary for the definition of the Impartial Observer theorem. Unlike Broome, Mongin concentrates his criticism on the distinction between preference judgments and utility amounts. It is important to note that Mongin (2001a) considers the causal interpretation of extended preferences made by Harsanyi (defined in sub-section 1.5.2) in order to make his criticism. In this interpretation "the process of identifying oneself with the other gives way to the process of deducing what one's own preference would be under certain ideal conditions" (Mongin 2001a, 15).

Mongin identifies another view of extended preferences, which he considers as the standard view: "individual i prefers to be j in social state x than to be k in social state y" (Mongin 2001a, 15).[42] This latter interpretation is problematic because if the impartial observer i identifies himself with $j's$ or $k's$ personality to make extended preference judgments, it cannot be the impartial observer who makes them. Thus the standard view of extended preferences leads to difficulties regarding the personal identity of the impartial observer.

[41] The criticism made by Broome (1993) is questioned by Kolm (1994), who says that Broome has misunderstood the concept of "fundamental preferences" in his document. It is not our aim here to go into this discussion or into the difference between extended preferences and fundamental preferences.
[42] j and k refer to individuals.

When Harsanyi formalized the extended utility function in his paper of 1977, he defined a vector (R_j) that contains all the objective causal variables required to explain the preferences of individuals. Thus the extended utility function of each individual is represented by the following equation: $V_i = V_i[A_j, R_j]$, which implies that the utility that individual i would give to the social situation A_j is determined by the vector R_j. This function V_i "must be the same for all individuals i" (Harsanyi 1977, 58).

According to Mongin, the proposition made by Harsanyi "does not deliver uniformity of extended preferences, only uniformity of the considered utility amounts". Thus Mongin takes it as an assumption that the statement $V_i[A_j, R_j] \geq V_i[A_k, R_k]$ "might well hold for all i without representing any observer's preferences" (Mongin 2001a, 20).[43] That is to say that individual i can obtain more utility in the social situation A_j with causal variables R_j than in the social situation A_k with causal variables R_k. This does not however indicate that someone prefers the first extended alternative to the second.

This conclusion allows Mongin to give a weaker version of the Impartial Observer theorem, which exclude the uniformity of extended preferences as a principle that must be held:

> Assume that for each individual, the extended preference relation satisfies the (VNM) axioms, and that (EC) and (CO) hold. Then, for each i, there exists a representation w_i of his moral preferences for X, such that: for all $x \in X$,
>
> $$W_i(A) = 1/n \sum_{j=1}^{n} U_{ij}(A), \qquad (7)$$
>
> where $U_{i1}, \ldots, U_{in}$ are utility representations of the individuals' actual preferences (Mongin 2001a, 22)

We observe that the implication that the extended preference relation does not satisfy the uniformity principle is that the utility representation of j's preference depends on the particular observer i. This hypothesis proposed by Mongin reveals that the subjective character of preference judgments may not be avoided even in the case of an impartial observer.

In brief, these two criticisms of the extended preferences proposition use different methodologies. They both reveal that subjective attitudes cannot be "objectified" when we make interpersonal comparisons of utility. Neither of these criticisms highlight however who takes the role of the impartial observer, because they are focused on demonstrating that values

[43] Our notation is not the same as that used by Mongin.

have to be part of the analysis but not on who makes these value judgments. With this aim in mind, we present the analysis of MacKay (1986), who identifies the impartial observer as the 'investigator'.[44] It is important to note that Mackay's approach uses the standard interpretation of extended preferences that Mongin discards when analyzing the impartial observer. We consider that even if this interpretation of extended preferences leads to "the puzzling difficulties of personal identity", its analysis can be useful in determining to what extent such difficulties may be insuperable (Mongin 2001a, 29).

1.5.4.3 Mackay's approach: The Impartial Observer's Identity

Mackay identifies three themes that are essential for him in Harsanyi's impartial observer framework: (1) 'the mental shoehorn maneuver' (empathy), (2) the counterfactual judgment, and (3) the logical basis of interpersonal comparisons of utility. The relationship between the first and third themes is one of the author's concerns, because when the 'investigator' imagines himself to be in the shoes of someone else, this depends on the satisfaction of certain conditions such as tastes, education level, social background, cultural values and psychological makeup (Mackay 1986, 318).

The question asked by Mackay is how far he can get the imagination of the investigator when he swaps places with another individual. In essence, Mackay doubts why anybody would think that "He" (in his role as an investigator) has any special authority about how "He" (with the objective and subjective attitudes of another individual and none of his own) would react in different circumstances? The importance of this proposition lies in the possibility that the economist cannot put aside his own objective and subjective attitudes at a specific moment in order to evaluate those of another individual.

The impartial moral point of view of the economist is therefore questioned because there is the risk of 'experimental contamination' which appears when the economist has to infer the psychological laws of another individual based on his own actual psychological relations. Mackay sheds light on this issue as follows:

> The "me" that appears under these extreme hypotheses -a person in his position, with his taste, his education, his social background, his cultural values, and his psychological makeup-would indeed display his reactions, but wouldn't be a "me" toward which I, the investigator, had any special epistemic authority. I, the

[44] This main objection also arises in Kaneko (1984).

> investigator, would be no better off asking questions about this "me" than I would about the original other person (Mackay 1986, 322).

According to Mackay, the economist is not able to take the role of the impartial observer because his values could be intermingled with those of the economic agent. The importance of this conclusion lies in two principles: (1) the rejection of the extended preferences proposition for making interpersonal comparisons of utility, and (2) the problem of experimental contamination, which the economist undergoes when he puts himself in the shoes of someone else.

This experimental contamination problem shows that economists cannot assume an 'epistemic authority' because they cannot express their own mental states and experiences. It is because the causal variables defined by Harsanyi are a black box and it will be difficult for the economist (as an impartial observer) to identify what his own values are and what the values of the economic agent are. Mackay's analysis reveals that even if an economist takes the role of impartial observer under Harsanyi's framework, he cannot be completely objective making ICU's. On the other hand, this analysis also reveals that the causal variables defined by Harsanyi keep Pandora's box that represents the different value judgments of economic agents closed.

Mackay's criticism might be perceived as a real difficulty when making the distinction between the value judgments of economists and those of economic agents due to experimental contamination. We think however that if it is assumed that an economist is the impartial observer, it is essential to know to what extent her value judgments may be useful to recommend a specific action for the benefit of general social interests. In order to overcome the difficulty posed by experimental contamination in making the distinction between the value judgments of economists and those of economic agents, Mongin's (2006b) framework about the neutrality value problem becomes enlightening. It means that the economist *qua economist* has the capacity to recognize when she makes a judgment of value or a judgment of fact.

The main problem of Harsanyi's framework is however the uniformity of extended preferences, which are based on causal variables. As we already know, these causal variables explain the subjective attitudes and personal preferences of individuals, which lead us to think that they form a set of motivating reasons as described by Dietrich and List (2013). Motivating reasons may be determined by normative reflections when the economic agent follows a conscious process of revising his or her preferences given a set of alternatives. As a result, the causal variables may be formed by both motivating reasons and normative reasons. So, we consider

that the classification of value judgments made by Sen (1967), that we already studied in this chapter, can be useful when studying the normative reasons that can justify the preferences of an economic agent.

1.6 Conclusions

Throughout this chapter we have shown the importance of the distinction between the value judgments of economists and those of economic agents in economics. The fact of making this distinction explicit allow us to understand, on the one hand, that it is important to consider the normative reasons that lead economic agents to justify their preferences, and on the other hand, that economists cannot avoid making value judgments in their professional work.

The study of authors such as Myrdal and Weber are essential in order to recognize their positions about the role of value judgments in social science, even if they seem totally opposite. They consider that value judgments are perceived differently depending on whether one has the role of a social scientist or the role of an economic agent. We were able to verify that both authors try to avoid biases in the study of social phenomena, but in a different way.

Mongin's framework and Sen's framework were important to examine the nature of the value judgments of economists and those of economic agents. Mongin's framework allowed us to better understand the debate of welfare economics about the value judgments of economists, and to recognize the different degrees of neutrality that economists can have in their professional work. Sen's framework allowed us to study a set of criteria that would be useful to determine the nature of value judgments of economic agents.

If Mongin's analysis seems fitting to study the role of value judgments of economists in their professional work, the same cannot be said of Sen's analysis to study how economic agents use their value judgments to justify their preferences. As a result, it is essential to continue studying how the set of valuations that economic agents hold can justify their preferences. One way to do that is to examine the set of valuations that a person can hold according to the different dimensions of her personality (person's plural identities), which has been analyzed by Sen (2006) and Binder (2019).

The case study about interpersonal comparisons of utility showed us how Harsanyi claims that ICU's are not based on ethical or political value judgments on the ground that an individual's

preferences are based on impersonal attitudes (extended preferences) guaranteed by the existence of an impartial observer. Extended preferences are however widely criticized for relying on causal variables. The three criticisms proposed in this chapter use different methodologies, but we consider that all three highlight the important role that value judgments play when one makes interpersonal comparisons of utility.

The criticism made by Mackay to Harsanyi's impartial observer framework allowed us to verify that it is necessary to have a theoretical framework to make the distinction between the value judgments of economists and those of economic agents in normative economics. For this purpose, two frameworks have been useful: Sen's classification of value judgments and Mongin's theses about value neutrality in economics. This implies that we consider that it is important to identify the role of an impartial observer. We supposed that economists take on this role because they can make judgments of values and judgments of facts about individuals' well-being with a certain degree of neutrality, while we assumed that the value judgments of economic agents justify their preferences, when making choices.

Chapter 2

One Way in which Social Choice Theory has Contributed to the Conception of Freedom

---

Abstract

This chapter takes freedom as a normative criterion for evaluating alternative states of affairs. We discuss how freedom is dealt with in social choice theory, especially in the context of preference-based approaches. We start with an analysis of the contributions of moral and political philosophy to the concept of freedom. This allows us to determine to what extent these contributions influence the definition of freedom and freedom of choice. Then we focus on examining the different ways to measure freedom in economics. Thus we study preference-based approaches, that arise in response to Pattanaik and Xu's cardinality rule, which states that the more options a person has, the bigger her freedom of choice is. Another way to answer to this cardinality rule is the diversity approaches, which consider the degree of similarity or dissimilarity between alternatives in a set. We revisit the origins of preference-based approaches in order to establish the tools they use from social choice theory to develop the conception of freedom in economics. Finally, we show that the importance of preferences, consequences, and notions such as incompleteness and maximality in social choice theory play an important role in the study of preference-based approaches.

2.1 Introduction

In the first chapter of this dissertation, we have studied the importance of making a distinction in economics between the value judgments of economists and those of economic agents. Now, we are going to study one particular value with the aim of establishing how economists study this value and how economic agents use this value to make their choices. We choose freedom because this value plays an important role in the economic, political, social, and cultural systems of Western societies. If we focus on the role of freedom in economics, it is quite clear that economists have shown interest in the studying of freedom of choice, free exchange, free entrepreneurship, free trade, free markets, etc. Freedom can be considered by economists as a normative criterion to evaluate alternative states of affairs.

In this chapter we inquire as to how the opportunity approaches (or preference-based approaches) use the tools of social choice theory to develop a conception of freedom in economics. This leads us to focus on the analysis of the freedom of choice that an individual has over an opportunity set at a given instant. This literature is concerned with the definition of individual freedom. We choose this field because preferences have been extensively studied in economics. Preference is an essential concept both in positive economics (to characterize rationality) and in normative economics (to evaluate choices and institutions in terms of welfare). There is a plurality of interpretations of preferences that can be used in different context, which is considered as a strength by authors like Sen (1996, 1997a). We think that this plurality of interpretations leads us to the development of three different ways to include preferences in the measurement of freedom of choice: the actual preference ordering of the person in question, all possible preference orderings that the person may have under uncertainty, and all preference orderings that the reasonable person may have (Dowding and van Hees 2009; van Hees and Wissenburg 1999; van Hees 2000).

In order to determine what tools preference-based approaches use from social choice theory to develop the conception of freedom, we consider that it is important to examine the origins of

these approaches. This will allow us to show: (1) How individual preferences in social choice theory influence preference-based approaches. (2) How consequences in social choice theory lead us to study two aspects when assessing freedom: the process aspect of freedom and the opportunity aspect of freedom (Sen 1990; 1993b; 2002). (3) How notions such as incompleteness and maximality that come from social choice theory are important in the analysis of multiple preferences and intersection rankings.

Opportunity approaches are one of the ways to responds to the cardinality rule proposed by Pattanaik and Xu (1990), which establishes that freedom of choice depends on the number of alternatives in a set. Another way to respond to this cardinality rule is to take into account the degree of similarity or dissimilarity between alternatives in a set. These approaches are known in the literature as diversity approaches. However, these proposals fail to take into account either the importance of the cardinality of a set, reducing freedom of choice to diversity, or the degree of distance between alternatives in a set, focusing only on the number of partitions in a set containing similar alternatives (Dowding and van Hees 2009; van Hees 2004).For this reason, in this chapter we will not analyze diversity approaches in the study of the measurement of freedom of choice.

All this considered, this chapter will have four sections. The first section presents the contribution of philosophers to the conception of freedom. This allows us to understand the most important debates with regard to the conception of freedom in moral and political philosophy. The second section analyzes the different ways freedom in economics can be measured. We present at first the distinction between freedom and freedom of choice, which can shed light on a different way to respond to the cardinality rule of Pattanaik and Xu (1990). Later on, we present the two traditional responses to the cardinality rule: diversity approaches and opportunity approaches. The third section studies the origins of preference-based approaches with the aim of establishing the tools they use from social choice theory. The fourth section concludes by making general remarks about the different ways to overcome the dilemma posed by Pattanaik and Xu (1990), the importance of the preference-based approach in the measurement of freedom, the appropriability of the preference-based approach to study individual freedom, and the most important tools that social choice theory has provided to the preference rankings for the evaluation of freedom.

2.2 Freedom in Moral and Political Philosophy

This section presents in a non-exhaustive way how freedom has been studied in philosophy. We focus on moral and political philosophy and ignore debates about free will and determinism The aim is to study the most important contributions of these domains to the conception of freedom, which will be useful to evaluate to what extent economists made use of these propositions or not in their study of freedom. We consider the seminal work of Berlin (1969 [2002]) who proposes two concepts of freedom (negative and positive freedom) that become two fields of study difficult to avoid when doing research on freedom. But some authors try to overcome this dichotomy, so they propose a third conception of freedom, such as Pettit (1997) and his republican conception of liberty, or a triadic relation, such as MacCallum (1967), just to mention a few contributions.[45]

2.2.1 The Negative and Positive Freedom Dichotomy

> Here I can only restate my original thesis that it seems to me patently inconsistent to assert, on the one hand , that all events are wholly determined to be what they are by other events (whatever the status of this proposition), and, on the other, that men are free to choose between at least two possible courses of action – free not merely in the sense of being able to do what they choose to do (and because they choose to do it), but in the sense of not being determined to choose what they choose by causes outside their control (Berlin 1969 [2002], 5).

There is no doubt that Isaiah Berlin is conscious about the importance of the debate between free will and determinism.[46] He affirms however, from the introduction of his seminal

[45] In this chapter we use the terms liberty and freedom as synonyms.

[46] Free will "has traditionally been conceived of as a kind of power to control one's choices and actions." (O'Connor and Franklin 2021). On the other hand, "the world is governed by (or is under the sway of) determinism if and only if, given a specified way things are at a time t, the way things go thereafter is fixed as a matter of natural law." (Hoefer 2016). In essence, the problem of free will and determinism lies in the difficulty of reconciling the idea that we have power to control our choices and actions with the idea that impersonal forces that we cannot control may causally determine our choices and actions (Caruso 2018).
As a result of this difficulty, philosophers are divided into two fields that seem irreconcilable. On the one hand, those philosophers that argue that free actions cannot be causally determined are part of the field known as incompatibilism. On the other hand, those philosophers that affirm that there is no incompatibility between free will and determinism are part of the field known as compatibilism. The solution to this problem has been characterized by an endless dispute between compatibilists and incompatibilists. For this reason, we are not going

contribution, *Liberty*, that his objective is not to discuss this metaphysical problem. One of his most important contributions is the study of the difference between positive and negative liberty. We study here these two conceptions of liberty and some debates around it. We focus on the paradox of positive liberty, republican liberty, and freedom as triadic relation.

Berlin says that negative liberty attempts to answer to the question "What is the area within which the subject – a person or group of persons – is or should be left to do or be what he is able to do or be, without interference by other persons?". Positive liberty attempts to answer the question "What, or who, is the source of control or interference that can determine someone to do, or be, this rather than that?" (Berlin 1969 [2002], 169). These questions capture the essence of the two conceptions of liberty. On the one hand, the absence of interference by other persons is crucial in the negative conception of liberty. A person is negatively free if no one is obstructing her doing whatever she could want to do. On the other hand, the presence of control on the part of the individual is crucial in the positive conception of liberty. An individual is positively free if she is able to control her own destiny in her own interests. We propose an example in order to better understand these two conceptions of liberty.

Imagine Mrs. Z is riding a bicycle through town, and she comes to a crossroads. She turns right, but no one was preventing her from going left or straight on. Next, she comes to a fork in the road. She turns left, but no one was forcing her to take one way or the other. There are no pedestrians or cars and there are no diversions. She seems, as a cyclist, to be completely free. But this situation might change if we consider that the reason Mrs. Z went right and then left is that she is alcoholic, and she is desperate to get to the liquor store before it closes. Rather than riding a bicycle, she feels she is being ridden, as her urge to drink leads her irrepressibly to turn the wheel first to the right and then to the left. Moreover, she is perfectly aware that her turning left at the fork means she will probably miss an appointment she cares about very much. She longs to be free from this irrational desire that is not only threatening her health but also stopping her from doing what she thinks she ought to be doing.

In the above story, Mrs. Z is free in the negative sense because there are no external obstacles stopping her to turn right and then left. It means that there is no interference of other human beings (e.g., pedestrians, drivers, etc.) with whatever she might want to do. In contrast, Mrs. Z

to take part in this debate, but we consider that this dispute highlighted the importance of the control that human agents may have over their actions (free will) in philosophy.

is unfree in the positive sense because she is moved by an irrational desire that she fails to control. She is an object of her addiction which is preventing her from doing what is her true interests.

2.2.2 The Paradox of Positive Liberty

Berlin is one of the clearest sources of two of the arguments that question positive liberty. First, it is suggested that the conception of positive liberty carries with it a danger of tyranny. Second, since an individual is the master of a situation, then she is freer when, instead of getting rid of restraints faced by her real desires, she is manipulated into giving up those desires.

Berlin justifies the tyranny argument with the idea of a divided self. To illustrate this, we take the alcoholic woman in our story, because Mrs. Z is both a self that desires to go to an appointment and a self that desires to go to the liquor store, and these two desires are in conflict. According to the framework proposed by Berlin (1969 [2002]), we suppose that the self that wants to get an appointment is a 'higher' self, and the self that is an alcoholic is a 'lower' self. The higher self is "identified with reason, with my 'higher nature', with the self which calculates and aims at what will satisfy it in the long run, with my 'real', or 'ideal', or 'autonomous' self, or with my self 'at its best'." In contrast, the lower self is "the pursuit of immediate pleasures, my 'empirical' or 'heteronomous' self, swept by every gust of desire and passion." (Berlin 1969 [2002], 179). In this way, an individual is free, when her higher self is in control, and she is not dominated by her passions or her empirical self. The consequence of this division is that some individuals are more rational than others and can coerce others in the name of their rational interests. That is to say, the more rational individuals can force less rational people to do rational things, in order to raise them to a higher level of freedom. In brief, a tyrant can bully, oppress, or even torture people in the name of their rational interests.

In order to better understand the second argument that is part of this paradox, we present Berlin's example of the wounded leg: "there are two methods of freeing myself from pain. One is to heal the wound. But if the cure is too difficult or uncertain, there is another method. I can get rid of the wound by cutting off my leg." (Berlin 1969 [2002], 182). If this person trains to deprive herself of what her leg normally allows her to do, then she shall not feel the lack it. This strategy makes the individual freer by retreating into "the inner citadel" in which the individual undergoes a process of self-transformation that makes her immune to any outside forces

(society, public opinion, etc.). However, even if this state can be achieved, Berlin considers that it is implausible that the strategy of retreating into an inner citadel can enlarge liberty. On the one hand, there is a risk of covering up important forms of oppression. On the other hand, the removal of desires may be a result of brainwashing by external forces, which is not compatible with any conception of liberty.

Both arguments constitute a strong criticism of the conception of positive liberty. John Christman proposes one of the responses to this criticism. He argues that the inner citadel argument can be overcome when an individual undergoes "a self-conscious and (let us say) rational program of character change," and some obstacle appears that would have stopped her from satisfying some previously discarded desire. This obstacle is then not a constraint for her, and it has no effect on her freedom (Christman 1991, 353). As an illustration, imagine a Buddhist sage who has spent the previous 5 years in the same room meditating and from which he will never desire to leave. If the door of the room is chained, then his freedom of action is not affected by these chains. This result is possible because if the desire changes of the Buddhist sage take place before the placing of the chains, then we cannot say that the chains tied to the door are the cause of the preference change.

With regard to the tyranny argument, Christman affirms that many critics of positive freedom and even some defenders of this concept have made the mistake of ignoring "the special and intrinsic value that is contained in self-government itself" and "the capacity for character formation and self-identity, which are themselves intrinsically valuable." (Christman 1991, 358). To consider then the actions of an individual as positively free, what matters is the way in which desires are formed, that is to say "the procedural conditions of autonomous preference formation that are constitutive of freedom," and not the "content" of an individual's desires (Christman 1991, 359). Take the example of the son of a family of Orthodox Catholics who claims to follow the doctrines of his family and the community in which he lives. According to Christman, this man is positively unfree if his desire to follow the doctrines of Orthodox Catholics was oppressively imposed by indoctrination, manipulation, or treachery. But this man is positively free if his desire was formed considering rationally other options. In this way, forcing him to do certain things can never make him freer, whereby the tyranny argument is avoided.

Berlin's distinction generates reactions like that of Christman who claims the importance of the positive conception of freedom, but also leads other theorists of liberty to go beyond the

negative conception of freedom. For these theorists, liberty is not just having a sphere of non-interference but also having certain conditions that guarantee non-interference. This is known as Republican liberty, or liberty as non-domination.

2.2.3 Republican Liberty

Republican liberty is considered as a third conception of freedom, which shows that the debate on freedom does not only focus on the dichotomy of positive and negative. Pettit (1997) argues that Berlin's conceptions of freedom allow a third possibility. Then, "This possibility would have one conceptual element in common with the negative conception – the focus on absence, not presence – and one element in common with the positive: the focus on mastery, not interference." (Pettit 1997, 22). The conception of freedom as non-mastery, or as Pettit prefers to say, non-domination, is distinct from the positive and negative conception of freedom.

Pettit affirms that it is easier to see the distinction between freedom as non-domination and positive freedom than the distinction between freedom as non-domination and negative freedom. "Non-domination by others is distinct from the positive ideal of self-mastery since the absence of mastery by others does not guarantee the achievement of self-mastery." (Pettit 1997, 22). The distinction between non-domination and non-interference by others may seem less obvious. Pettit shows that domination and interference are different 'evils.' A person may suffer domination without interference or undergo interference without being dominated.[47]

To illustrate the case of domination without interference, we take a slave who serves a kindly master. Even if the master does not interfere with the choices of his slave, the slave suffers from domination to the extent that she has a master. The slave enjoys non-interference to the extent that her master fails to interfere. Domination occurs because the master has the capacity to interfere arbitrarily, even if the master is never going to do so. With respect to the case of interference without domination, we take a policeman who asks a person to evacuate the train station due to the presence of a suspicious package. He does not allow her to take a train to go home. This interference is justified because he is acting in the interests of this person, whose life may be in danger. Here it is not possible to see the interference as an exercise of domination but it is not arbitrary either. "Interference can occur without domination, because interference

[47] Differences between republican freedom and negative freedom are debated. We do not focus on this debate because it extends well beyond the topic of this chapter. Carter (2019) proposes an introduction to this debate in his entry in The Stanford Encyclopedia of Philosophy.

need not involve the exercise of a capacity for arbitrary interference, only the exercise of a much more constrained ability." (Pettit 1997, 23).

Pettit proposes two reasons for the association of the conception of liberty as non-domination with the view of liberty in the republican tradition. First, in the republican tradition liberty is considered in terms of the opposition between citizen and slave. The condition of liberty is given by the status of someone who is not at the mercy of the arbitrary power of anyone else. This opposition between slavery and freedom is partly based on the experience of the Roman Republic. Second, as we already mentioned, under the republican conception of liberty, interference takes place without any loss of liberty. This interference is not arbitrary and comes from "the law and government that obtains in a well-ordered republic." (Ibid, 31). In the republican tradition, freedom is considered as a status that exists only in an appropriate legal regime. "As the laws create the authority that rulers enjoy, so the laws create the freedom that citizens share." (Ibid, 36).[48] According to the republican view, the laws respect people's common interests and they are not the tools of an individual's arbitrary will.

The non-domination must be in the presence of other people, not in isolation from the society. One of the strategies for achieving non-domination proposed by Pettit, consists in introducing a constitutional authority (e.g., an elective agent) that will not allow other parties to have the power to make arbitrary interferences nor the power of punishing that kind of interference. In this way, this strategy leads to end domination because the actions of the constitutional authority take into account the interests and ideas of citizens. Pettit says that "non-domination is itself a form of power. It represents a control that a person enjoys in relation to their own destiny and such control constitutes […] the power of the agent who can prevent various ills happening to them." (Ibid, 69).

2.2.4 A Triadic Freedom-Relation

> This distinction between positive and negative freedom has, however, stood in the way of this approach. It has encouraged us to see differences in accounts of freedom as resulting from differences in concepts of freedom. This in turn has encouraged the wrong sorts of questions. We have been tempted to ask such questions as "Well,

[48] Skinner (2002 [2006]) also studies the republican conception of freedom. "To be free as a citizen, therefore, requires that the actions of the state should reflect the will of all its citizens, for otherwise the excluded will remain dependent on those whose wills move the state to act." (Skinner 2002 [2006], 252) .

> who is right? Whose concept of freedom is the correct one?" or "Which kind of freedom do we really want after all?" (MacCallum 1967, 320).

MacCallum thinks that the dichotomy between positive and negative freedom has led us to the wrong questions. He proposes only one basic concept of freedom, on which the positive and negative conceptions of freedom converge. He defines freedom as a triadic relation, a relation between three terms: an agent or agents, determined constraints or preventing conditions, and determined doings or becomings of the agent. An agent (x) is free (unfree) from a constraint (y) to do (not do, become, not become) an action or condition of character (z). Therefore, any claim about freedom or unfreedom in a given situation has to make suppositions about what counts as an agent, what is to count as a constraint or restriction on freedom, and what matters as a purpose that the agent can be designated as either free or unfree to accomplish.

In order to illustrate MacCallum's triadic relation, we again consider the example of the woman riding a bicycle to the liquor store. If we say that the woman is free, it probably means that an agent, consisting in the woman's empirical self, is free from external obstacles (e.g., pedestrians, drivers, etc.) to do whatever she might want to do. If, in contrast, we say that the woman is unfree, it probably means that an agent, consisting in the woman's rational self, is unfree by internal, psychological restrictions to accomplish some rational plan (go to an appointment, instead of going to the liquor store). It is important to note that in both assertions there is a negative and a positive part. It means that each assertion about freedom supposes both that freedom is freedom from something and that it is freedom to do or become something. The dichotomy between negative and positive freedom is a false one. It is deceiving to say that when we assume the woman as free, we use the negative conception of freedom; and when we assume the woman as unfree, we use the positive conception of freedom.

According to MacCallum the difference between negative and positive freedom is given by the way in which each of the three variables in the triadic freedom-relation is interpreted. It means that each camp identifies in a different way "what can serve as agent, preventing condition, and action or state of character vis-à-vis issues of freedom." (MacCallum 1967, 327).

We now present how adherents of negative and positive freedom treat the three variables in the triadic freedom relation. With respect to the agent, adherents of negative freedom conceive the agent as an individual human being who has empirical beliefs and desires. In contrast, adherents of positive freedom conceive the agent when there is a contraction of his limits. It means that the agent is identified as the ""real" or the "rational" or the "moral" person who is sometimes hidden within, or has his seed contained within, the living human body." (MacCallum 1967,

324). Sometimes, however, they conceive the agent when there is an expansion of his limits. It means that the agent is identified as a person whose desires and aims are in relation to some collectivity to which she belongs.

With regard to preventing condition, negative theorists count something as a preventing condition when it is the result of arrangements made by human beings, which represent external obstacles as restrictions on freedom. Positive theorists generally have a wider view of what counts as a preventing condition than negative theorists. They consider that a person may also be constrained by internal obstacles, such as ignorance, irrational desires, or illusions.

With reference to what counts as a purpose a person can be free to have, adherents of negative freedom tend to have a wider view of this variable than positive theorists. They consider that the agent may desire any set of actions or states. Adherents of positive freedom tend to limit the relevant set of actions or states those that are rational, real, or moral.

Note that MacCallum's conception of freedom has no place for a dichotomy between positive and negative freedom. There are rather different interpretations of the triadic-freedom relation. MacCallum considers that this dichotomy has led to assigning writers on freedom to a category to which they do not really belong.[49] He concludes that establishing the basis of the debate on freedom as two irreconcilable fields can lead to the distortion of the conceptions of freedom.

On the whole, we have studied the debates that arose from the dichotomy between negative and positive freedom proposed by Berlin. We observed that from this dichotomy arose not only strong defenders of one field or the other, but also new alternatives that allowed us to go beyond a dyadic relation. This is the case of the Republican liberty or liberty as non-domination and the triadic freedom-relation. In the next section, we study if economists consider the debates already mentioned in their studies on the conception of freedom in normative economics.

2.3 Exploring Ways to Measure Freedom in Economics

In the middle of the nineteenth century, freedom was considered as an important criterion to compare alternatives of economic policy. In his essay *On Liberty,* John S. Mill analyzes

[49] MacCallum takes as an example Locke, who is considered to be in the field of classical liberalism and therefore of negative freedom. However, he also says that liberty is not to be confused with 'license', and that it is possible to restrict actions that are in the agent's own interests. The following question arises: "Why doesn't all this put him in the camp of "positive" freedom (…)?" (MacCallum 1967, 322).

different cases showing how freedom can be useful to evaluate alternative policies, e.g., "the State should take measures to render the means of drunkenness more costly, or add to the difficulty of procuring them, by limiting the number of the places of sale" (Mill 1859 [1998], 119). Mill states that to render alcoholic beverages costlier, the State can apply a tax. The taxation can be justified by the State because alcohol produces a big amount of revenue. In contrast, the limitation of the number of places to sell is more difficult to be justified by the State because it may involve treating part of the population as children or savages. This example shows us that the two measures restrain the agent from the consumption of alcoholic beverages. To know the best measure to take, Mill uses freedom as a criterion to compare these alternatives.

Considering freedom as a normative criterion has not been a matter on which economists have reached consensus. At the end of the nineteenth century, the marginalist revolution led to the focus on individual subjective preferences as the natural criterion to evaluate the social goodness of any social state. This approach known as welfarism dominated the debate in normative economics.[50] But even if individual subjective preferences have been the dominant criterion to assess social welfare, other criteria have also been studied "The task of normative economics is to investigate methods and criteria for evaluating the relative desirability of economic states of affairs" (Mongin 2002, 145).

We can identify the areas that have conceived freedom as a normative criterion in economics. The first area is based on the theory of rights proposed by Sen (1970) in his article on the Paretian liberal paradox and concerns with the mutual compatibility of individual freedoms. This paradox is based on the inconsistency of three conditions: unrestricted domain, the weak Pareto principle, and the condition of minimal liberty. There is a large body of literature devoted to either dropping at least one of the conditions or presenting a weaker version of at least one of the conditions.[51]

The second area studies the fairest way to distribute individual freedoms. This area is rooted in philosophy, particularly in Rawls's (1971 [1999]) seminal contribution.[52] The problem studied by this literature is basically the following:

[50] See Mongin (2002) who studies the progress in normative economics. See also Baujard (2017) who studies the historical evolution of welfare economics.
[51] See Sen (1976; 1983) for his position on the different reactions generated by the impossibility of the Paretian liberal.
[52] Rawls considers that the equal basic liberties in his first principle of justice are: "freedom of thought and liberty of conscience; the political liberties and freedom of association, as well as the freedoms specified by the liberty and integrity of the person; and finally, the rights and liberties covered by the rule of law." (Rawls 1981, 5).

> Given a notion of freedom embodied in a particular ranking of opportunity sets, a natural egalitarian and freedom respectful normative objective would seem to consist in equalizing as much as desired the individual opportunity sets and, once this level of equalization is achieved, in letting the individuals choose from their opportunity sets their best option according to whatever objective they may pursue (Laslier et al. 1998, 4).

However, the appropriate metric of equalization of opportunity sets has been an important subject of debate because individuals are different with respect to their state of health, productive capacities, manual dexterity, etc.

Finally, the third area is based on the freedom of choice that an individual has over an opportunity set at a given point in time. This area is concerned with the definition of individual freedom. The debate is about defining and measuring individual freedom of choice. It is focused on the "objectivist" and "subjectivist" conceptions of freedom of choice. The "objectivist" conception of freedom of choice refers to "a ranking of opportunity sets based upon the physical attributes of the set itself and not upon the immediate subjective preferences of the chooser." The "subjectivist" conception of freedom of choice refers to the way by which "the immediate subjective preferences of the chooser" underlie the ranking of opportunity sets (Gravel, Laslier, and Trannoy 1998, 77).[53]

The debate between the 'objectivist' and 'subjectivist' conception of freedom of choice is of our interest because it allows us to analyze the difficulty that arises when defining and measuring individual freedom of choice. This debate particularly sheds light on whether individual preferences can contribute to individual freedom of choice. We consider it is as important to study the distinction between "freedom" and "freedom of choice" proposed by Carter (2004).

2.3.1 Freedom and Freedom of Choice

This distinction shows how normative economists and social choice theorists have often considered freedom simply as a shorter term than freedom of choice. Carter (2004) assumes that "a person has freedom iff she lacks constraints on the performance of an action (or set of actions)," and that "a person has freedom of choice iff she lacks constraints on the reasoned

[53] "An opportunity set is interpreted as the set of all options (commodity bundles, political candidates, lawful activities, etc.) to which an individual has access and from which she will, in some latter stage, make a choice" (Ibid, 77).

selection and performance of one or more of the items on an action menu" (Carter 2004, 68–9).[54] The difference between the two concepts is the term "choice," which implies the possibility of making a reasoned selection.

Carter is inspired by the triadic freedom-relation proposed by MacCallum (1967). He assumes that both freedom and freedom of choice are triadic relations. The former, between an agent, restrictions, and the performance of an action; the latter, between an agent, restrictions, and the selection of an action from an action menu. Relying on a triadic relation to make the distinction between freedom and freedom of choice allows him to identify the third variable capturing the essential difference between the two terms. "In the case of 'freedom,' the third variable is an action (or set of actions); in the case of freedom of choice, the third variable is the reasoned selection of an action (of one's own) from an action menu—that is, from a range of at least two alternative actions" (Carter 2004, 71–2).

Carter's distinction between freedom and freedom of choice serves as a tool to analyze the seminal work proposed by Pattanaik and Xu (1990). The cardinality rule suggested by the two authors through three axioms is a pioneering proposition for evaluating the degrees of freedom (of choice?) that different opportunity sets can offer. According to this rule one opportunity set is ranked higher than another in terms of freedom (of choice?) if and only if it has a larger number of items.[55] Carter aims to assess each of the three axioms with regard freedom, on the one hand, and freedom of choice, on the other.

Pattanaik and Xu deduce the cardinality rule from the following three axioms:

Axiom 1. Indifference between No-choice Situations (INS)

Two sets that contain only one item always offer the same degree of freedom. "If neither of two feasible sets offers the agent any choice in the sense that both the feasible sets are singletons, then the degrees of freedom offered by the two feasible sets are identical" (Pattanaik and Xu 1990, 386–87).

[54] Carter is interested in both the opportunity concept of freedom and choice, and not in the exercise concept of freedom and choice. When freedom is considered as an opportunity concept, it means the performance by the agent of some actions or actions given a lack of constraints. By contrast, when freedom is considered as an exercise concept, "it means the performance by the agent of some action or actions." (Carter 2004, 64). In the case of choice, a choice the agent *has* denotes an opportunity concept. On the other hand, a choice the agent *makes* denotes an exercise concept (Ibid, 66).

[55] They also say in the same paragraph that "under these axioms, judgements about *the degree of freedom* have to be based exclusively on the number of options available in the choice situation under consideration." (Pattanaik and Xu 1990, 383). We observe that the distinction between freedom and freedom of choice is not entirely clear in the seminal work of Pattanaik and Xu.

Axiom 2. Strict Monotonicity (SM)

Any set containing two alternatives (x, y) gives a person more freedom than having a set with only one alternative (x). This axiom "essentially embodies the principle that, in terms of freedom, a situation where the agent has some choice is better than a situation where the agent has no choice" (Ibid, 387).

Axiom 3. Independence (IND)

If you add or substrate the same element to or from two different sets, the freedom ranking of the two sets should not be affected. "IND requires that if A and B are two possible available sets and if x does not belong either to A or to B, then the ranking of A and B in terms of freedom corresponds to the ranking of $A \cup \{x\}$ and $B \cup \{x\}$" (Ibid, 387).

The cardinality rule deduced from these axioms is problematic even for Pattanaik and Xu, who suggest in their article by an example that IND is not as plausible an axiom as the first two axioms. They suppose alternative modes of transport: {train} and {blue car}. According to INS, these two modes of transport offer the same degree of freedom {train}$\sim${blue car}. Suppose we add another alternative (red car) to both sets. Then we have {train, red car}$\succ${blue car, red car}. This way of reasoning violates IND because when adding the same element (red car) to the sets {train} and {blue car}, the freedom offered by the two sets should be ranked equally. In other words, the independence axiom does not take account of the degree of dissimilarity between the alternatives. They conclude their paper: "our formal structure itself does not contain any information about closeness or similarity of different alternatives" (Ibid, 390).

We now study Carter's analysis in order to solve the failure of IND. He reconsiders the three axioms of Pattanaik and Xu and gets interesting results.

First, he concludes that INS is obeyed by freedom of choice and not by freedom. He argues that it is surely indisputable that if no-choice situations offer 'no choice', then they offer 'no freedom of choice'. But INS can be challenged as an axiom of freedom. It depends on the way a constraint of freedom is defined. He argues this as follows: if a coercive threat lead a person to do x, then she does not have the option not to do x. She does not have the freedom of choice between x and *not-*x, but the availability of x gives her some freedom.

Second, Carter assumes that IND is obeyed by freedom but not by freedom of choice. He affirms that "an agent can enjoy an increase in the number of actions available to her (and thus an increase in her freedom) without enjoying an increase in the number of actions from which

she can make a reasoned selection." (Carter 2004, 77). To illustrate this point, we take the example of an agent whose set of options increases from including the drinking of a can of coca to including the drinking of one or the other of two cans of coca that differ only by their barcode. In essence, the increase in the number of cans of coca increases the freedom of this agent. Since there is no significant difference between the cans of coca, it cannot be said that the freedom of choice increased.

Third, Carter initially suggests that SM is obeyed by freedom and not by freedom of choice, for the same reasons that were given regarding IND. He proposes a way to "save" the monotonicity axiom in the case of freedom of choice. By reconsidering the interpretation of the term "new option." "The addition of a "new option" to an already existing set is necessarily the addition of a significantly different option." (Carter 2004, 77). This leads him to conclude that SM is obeyed by both freedom and freedom of choice.

Carter establishes that each one of the axioms is true by distinguishing between freedom and freedom of choice. INS is true of freedom of choice but not of freedom, SM is true of both freedom and freedom of choice, and IND is true of freedom but not of freedom of choice. We consider that Carter proposes a third way to overcome the failure of the independent axiom. It means that there are two other ways to overcome the dilemma introduced by Pattanaik and Xu (1990). One focuses on the opportunity aspect of freedom, and the other on incorporating information about the diversity of alternatives in the cardinality rule. Carter suggests that those focusing on the opportunity aspect of freedom are "thinking, at least implicitly, in terms of the concept of freedom of choice", and that those focusing on the diversity of alternatives are "thinking, at least implicitly, in terms of the concept of freedom […]" (Carter 2004, 80).

What allows Carter to suggest that the group of theorists that studies the opportunity aspect of freedom may focus on the concept of freedom of choice, is that in general these approaches can be considered resulting from criticisms made to INS and SM. On the other hand, when Carter affirms that the group of theorists that studies the diversity approach may focus on the concept of freedom, is that this approach concentrates on redefining IND.

We are going to present the most important elements of these two responses to the cardinality rule in order to analyze the differences with Carter's proposition, and also to determine to what extent the solutions proposed by these two alternatives contribute to overcoming the dilemma proposed by Pattanaik and Xu (1990), according to which the independence axiom does not consider the degree of dissimilarity between many alternatives.

2.3.2 Freedom and Diversity

The diversity or non-preference based approach deals with the question of how to consider both the information about the cardinality of options in an opportunity set and the degree of distance among them (van Hees 2004).[56] Pattanaik and Xu (2000) propose an approach which takes into account the similarity or dissimilarity of alternatives. Their model is based on the cardinality of the smallest similarity-based partition of an opportunity set. The larger the number of partitions in an opportunity set, the more diversity the opportunity set offers. In order to illustrate this point, let us once again take their example on modes of transport: {train} and {blue car}. The opportunity set {train, blue car} can be split up in two subsets {train} and {blue car}, since a train and a car are dissimilar. However, since a blue car is similar to a red car, the number of smallest similarity-based partition is one {blue car, red car} (the set itself). For this reason, the opportunity set {train, blue car} offers more freedom of choice than {blue car, red car}.

Although this approach allows them to axiomatize an ordering focused on the cardinality of dissimilar options, it is still not satisfactory, since it does not take account of the degree of difference between dissimilar options. Such differences are important: a blue bus differs less from a red car than a red car does from a glass of red wine. Then the same problem of the original cardinality rule appears: the set {blue bus, red car} offers the same degree of freedom of choice as {blue bus, red wine}(van Hees 2004; Dowding and van Hees 2009).

Hence, others have chosen to take account of the degree of dissimilarity between options in a set. Rosenbaum (2000), conceptualizes freedom as follows:

> A function of the range of possible choices in terms of certain (at most n) characteristics, where each dimension i in $\Re^n$ represents one characteristic of the choice option, and where the coordinates $x_1, \ldots, x_n$ of a point in the space $\Re^n$ are to be understood as measures of the characteristics of a choice option (Rosenbaum 2000, 215–16).

The normalized maximum distance between a couple of alternatives in an opportunity set in the n-dimensional real space $\Re^n$, indicates the freedom of choice the opportunity set offers.[57]

[56] "The diversity issue is usually addressed without recourse to preferences." However, "the diversity issue might be conjoined with the opportunity issue." (Dowding and van Hees 2009, 375,378). Binder (2019) explores one way to connect the diversity and opportunity approaches. She suggests "a way to reveal information about the relevant differences between alternatives from preference information." (Binder 2019, 47).

[57] Klemisch-Ahlert (1993) proposes a similar approach. She also assumes that the alternatives of an opportunity set can be described as points in n-dimensional real space $\Re^n$. But she focuses on the convex hull of a set A with

Although this approach allows him to take account of the degree of difference between dissimilar alternatives, it fails to consider the number of alternatives a set includes. In order to illustrate why this is problematic, we take the example of the parties competing in a country's election. Suppose a situation in which the extreme left-wing party competes with the extreme right-wing party for the vote of people will provide the same degree of diversity as a situation in which, in addition to the two extreme parties, a center party is standing for election (Binder 2019).

In this way, we observe that these proposals that try to solve the diversity problem either fail to take into account the importance of the cardinality of a set, reducing freedom of choice to diversity, or they do not take account of the degree of distance between alternatives in a set, focusing only on the number of partitions in a set containing similar alternatives. Despite these results, Dowding and van Hees (2009) underline that an increasing literature on the measurement of (biological) diversity could help to find new ways to solve this problem (e.g. Nehring and Puppe (2002, 2009), Weitzman (1992), among others). Weitzman's contribution (1992) is a seminal work in the biological diversity literature. He takes as a given a cardinal dissimilarity metric between a pair of alternatives, in order to define "the marginal diversity an alternative adds to a set with the distance between the added alternative and the one in the set that is closest to it." (Binder 2019, 47).

In summary, to reformulate the IND axiom, it is necessary to take into account both the information about the degree of dissimilarity between alternatives and the information about the degree of dissimilarity between alternatives and sets of alternatives. However, "it is not altogether clear how such distances should be determined; in the literature one can find several ways of extending a 'normal' distance function to a function that also measures distances between sets and alternatives" (van Hees 2004, 256). For this reason, we do not focus on the analysis of this literature to the study of the measurement of freedom of choice. In contrast, we are going to focus on the preference-based approaches that at least have the same basic structure. In each of the cases studied in the next section, the original axioms of the cardinality rule are re-defined by taking account of preferences. However, the sets of relevant preferences are different. They may made up only of actual preference ordering of the person in question,

the aim of capturing "the idea of 'closeness' of an alternative x to the set A or the "similarity" of x to the alternatives in A"(Klemisch-Ahlert 1993, 196). The freedom of choice of a set of alternatives depends on the convex hull of that set: the larger the convex hull the more freedom of choice the set provides (Dowding and van Hees 2009).

or of all possible preference orderings that the person may have under uncertainty, or of all preference orderings that the reasonable person may have (van Hees 2000, 115).

2.3.3 Freedom and Preferences

Sen (1990) is the first author that has criticized the cardinality rule proposed by Pattanaik and Xu.[58] He considers that an agent's freedom of choice cannot be determined only by counting the number of options open to him but is also about the opportunities that these options provide him; that is, an agent's preferences over options are essential to assess the degree of freedom of this agent. The opportunity aspect of freedom is defined as follows:

> Freedom gives us the opportunity to achieve our objectives – things that we have reason to value. The opportunity aspect of freedom is, thus, concerned with our actual capability to achieve. It relates to the real opportunities we have of achieving things that we can and do value (no matter what the process is through which that achievement comes about) (Sen 1993b, 522).

According to Sen, the cardinality rule ignores the opportunity aspect of freedom. He argues that the problem lies in the INS and SM axioms, and not in the IND axiom as Pattanaik and Xu had affirmed in their seminal work.

With respect to INS, Sen (1990) gives the following example: suppose an individual that goes home from office, and she has to compare two ways in which she is forced to commute: (1) She is forced to walk in the usual way, or (2) She is forced to hop home on one leg. Sen considers that it is absurd to say that the individual has as much freedom in the second case as in the first, because it is obvious that any individual prefers to walk home normally.

In order to avoid this problem, Sen suggests the following reformulation of INS:

Let $x\ P\ y$ represent a binary relation of 'strict preference'. The relation P^* represents the relation 'offers strictly more freedom'.

For all x, y in X: if $x\ P\ y$, then $\{x\}P^*\{y\}$ (Sen 1991, 24).

[58] Pattanaik and Xu (1990) formalize a rule to judge the degrees of freedom of choice with three axioms. However, freedom of choice had already been studied by economists such as Sen (1988), Kornai (1988), among others.

It means that if x is strictly preferred to y, then the opportunity unit set of x offers strictly more freedom than the opportunity unit set of y.

With respect to SM, Sen (1990, 1991) argues that the addition of an alternative that an individual may consider as very unattractive, like the option of being beheaded at dawn, do not necessarily imply a strict enhancement of freedom. The addition of an alternative only leads to a strict enhancement of freedom if that alternative is strictly preferred to the alternative which an individual could already choose. In the following quote we can see Sen's formal reasoning for reformulating Pattanaik and Xu's SM axiom:

> If x is preferred to y as an alternative, it is possible to have a ranking of effective freedom that puts the unit set $\{x\}$ above the unit set $\{y\}$, and also puts the pair $\{x, y\}$ above each of the unit sets. We then have: $\{x, y\}P\{x\}P\{y\}$. The first part of the claim corresponds to Pattanaik and Xu's "Strict monotonicity" axiom […] it is possible to argue for weakening the first strict ranking P into a weak ranking R, if the alternative y is very like x […], or totally uninteresting as an option (Sen 1990, 471).

Sen (1990, 1991, 1993b) proposes to measure freedom in terms of a person's actual preferences. He is however aware of the importance of a person's future preferences, which are often uncertain. Considering a person's future preferences in the measuring of freedom is another way to determine the opportunity aspect of a person's freedom. Opportunity is about what a person can do not only now but also in the future. This way of determining the opportunity aspect of freedom takes into account the conception of preference for flexibility introduced by Kreps (1979). Kreps gives the following example to explain the concept of preference for flexibility: suppose that a person decides to make a reservation at a restaurant. She will choose the restaurant according to the menu of meals it will serve. It is assumed that she knows the menus at all restaurants that she might select. In the end, she will choose a meal, but her preliminary choice is a menu from which she will later choose her meal. We then take the case of a person that prefers a menu containing only fish to one containing only steak. But she strictly prefers a menu with both fish and steak to either of the first two, because it gives her greater flexibility. In this way, we can say that a decision maker prefers any set of opportunities to any of its subsets because of the greater flexibility associated with a larger set, in the presence of uncertainty about her own future preferences (Barberà and Grodal 2011).

An interesting approach showing that the opportunity aspect of freedom coincides with the notion of preference for flexibility is developed by Puppe (1996). He considers that a basic model in the context of the opportunity aspect of freedom is a two-stage decision model, "where

in the first stage a non-empty opportunity set (or menu) is chosen from which, in a second stage, exactly one alternative is chosen as the final outcome." (Puppe 1996, 175). Puppe aims to determine how the first-stage choices among non-empty opportunity sets are to be modelled. In order to do that, he proposes the following axiom:

Axiom F (Preference for Freedom of Choice). For all $A \in Z$: there exists $x \in A$ such that $A \succ A \setminus \{x\}$.

According to Puppe the intuition behind this axiom is as follows. Let A be the opportunity set of a decision maker in a specific choice situation. Assume that A includes all the alternatives that can possibly be chosen by a decision maker. Let x be one of the alternatives that there must be in A which the decision maker seriously takes into account for later choice. The decision maker considers x as essential in $\{x, y\}$ if, and only if, x is weakly preferred to y. A is then ranked strictly above $A \setminus \{x\}$ with regard to freedom of choice. Once Puppe has presented this axiom, he proposes a general model:

Let $\succcurlyeq$ be a freedom relation. A menu A dominates B if, and only if, A offers at least as much freedom as the union of A and B: $A \succcurlyeq A \cup B$. Then, Puppe indicates that axiom F together with an axiom of monotonicity and an axiom of independence of non-essential alternatives imply that A dominates B if and only if all essential alternatives of $A \cup B$ are available in A.[59]

With regard to uncertainty, Puppe assumes that if the decision maker strictly prefers x to y at the first stage of the two-stage decision model, then the decision maker will also strictly prefer x to y at the second stage. However, if the decision maker is indifferent between x and y in the first stage, then she will remain indifferent in the second stage. That is, the decision maker is uncertain about her indifference or incomparability judgments. Furthermore, assuming that the decision maker considers x as essential in $\{x\ y\}$ if, and only if, x is weakly preferred to y, then reference for freedom of choice is equivalent to the notion of preference for flexibility. Puppe

[59] Axiom of monotonicity: For all $A, B \in Z, B \subseteq A \implies A \succcurlyeq B$. So, if B is a subset of A then A gives at least as much freedom as B (Puppe 1996, 178).
Axiom of independence of non-essential alternatives: For all $A \in Z, E(A) \sim A$. Thus the assessment of the freedom given by non-empty menu A relies only on a non-empty subset $E(A)$ of essential alternatives (Puppe 1996, 182). It should be clear that "an alternative $x \in A$ will be called essential in A if $x \in E(A)$. If $x \notin E(A)$, x is called non-essential in A" (Ibid, 181).

concludes that "a menu A dominates another menu B in terms of freedom of choice if and only if A dominates B in terms of flexibility" (Puppe 1996, 177).[60]

Arrow (1995) makes an important contribution to the notion of flexibility. He argues that this notion is useful in the measurement of the opportunity aspect of freedom. He proposes to use "the concept of flexibility with respect to a probability distribution over preferences as the definition of freedom." (Arrow 1995, 11). It means that Arrow represents the uncertainty of individuals by a probability distribution over possible utility functions. The freedom offered by an opportunity set is then measured according to the expected indirect utility of an opportunity set.

Arrow says the set of preferences and the uncertainty about them can be explained in a number of ways. One interpretation "is from the point of view of the individual, at a stage where his or her preferences for the future are yet to be formed." Another interpretation "is from the point of view of a governing authority or, perhaps better put, a social compact. Here, we suppose that freedom to choose must be given in an impersonal way to individuals with a wide variety of tastes." (Arrow 1995, 11). In essence, Arrow proposes an approach based on multiple preferences, where these preferences represent the set of preferences that each individual could have in the future (potential preferences).[61]

In addition to this conception of potential preferences based in Arrow's approach, Sugden (1998) proposes two alternatives in order to measure opportunity by considering potential preferences. First, an opportunity set is valuable insofar as it includes options that might reasonably be chosen by a person. It means that for a given person, each of the options of an opportunity set corresponds "with a different but equally valid conception of her good." The second alternative considers a sociological interpretation of potential preferences. That is, the set of potential preferences consists of all those actual preferences that a person whose freedom is being assessed shares with other people with similar sociological characteristics (Sugden 1998, 325–26).[62]

[60] A menu A dominates another menu B regarding its flexibility if and only if all $s \in S$, $U(A,s) \geq U(B,s)$. Let S be a finite state space. Let $U(A,s)$ and $U(B,s)$ denote the derived indirect utility of menu A in state s, and of menu B in state s (Ibid, 189).

[61] See Foster (1993) who proposes a different approach based on multiple preference orderings.

[62] Note that Sugden (2003) assumes a critical view with regard the two alternatives to measure opportunity that he proposed in Sugden (1998). Even in Sugden (2010), he affirms that "I now think, opportunity is an open-ended concept: often, we cannot specify in concrete terms what a person does or does not have the opportunity to do, or what the value is of the things that she might do" (Sugden 2010, 48). Despite this, given that we are interested to

Pattanaik and Xu (1998) criticize the actual preference approach and the uncertainty preference approach because they fail to consider what they call the intrinsic value of freedom of choice. They follow the libertarian tradition that considers that "for a person to live a meaningful life, he must shape his life by making choices for himself" (Pattanaik and Xu 1998, 178) . For this reason, in order to assess the intrinsic value of freedom of choice, "the preferences that are crucial in such assessment are […] the preference orderings that a reasonable person in the agent's situation can possibly have" (Pattanaik and Xu 1998, 179–80).

In order to illustrate this way of measuring opportunity freedom, Pattanaik and Xu (1998) give the example of a woman who is totally convinced that she will never join the army. If women are banned from joining the army, these authors assume that a reasonable person in her position would feel that such a ban reduces women's freedom. The reason for this assumption is that any reasonable person in the woman's position considers that "given the woman's situation, she could have reasonably chosen to join the army […] even though she actually does not do so and even though she attaches zero probability to her wanting to do so." (Pattanaik and Xu 1998, 179). Agent's degree of freedom is a function of the preferences of a reasonable person in the agent's position. The preferences that a reasonable person may have play an important role in assessing the intrinsic value of freedom, and that this value lies in having choices between significant options. For this reason, not every option added to an opportunity set entails an increase in freedom of choice; nor does every reduction entails a decrease.

The study of Pattanaik and Xu (1998) is based on the paper of Jones and Sugden (1982), who consider the importance of reasonable preferences in the measurement of the opportunity aspect of freedom:

> A choice set has greater significance the more options it contains that could reasonably be chosen, and the greater the strength of preference that a reasonable person could have for the option he chooses over the options he might have chosen (Jones and Sugden 1982, 61).

Jones and Sugden consider that if only one option could reasonably be chosen, the opportunity set is not meaningful; and also, that if there is a number of options that could reasonably be chosen but no reasonable person would prefer one of them, then the opportunity set is not meaningful.

show the different conceptions of potential preferences, it seems important to us to take into account the two alternatives suggested by Sugden (1998).

The idea of using reasonable preferences for measuring the freedom an opportunity set provides, has however its limitations. We do not present an exhaustive list of the limitations of this approach; we only present some of them with the aim of showing the difficulty in constituting a criterion to judge the reasonableness of a person.

One problem with these approaches is that, "as long as the alternatives between which an agent can choose do not fully specify the agent's situation, it will always be possible to adjust the circumstances in such a way as to 'fit' a preference ordering" (van Hees and Wissenburg 1999, 75). Consider, for instance, an individual who has two options, a check for one million euros (x), and be beheaded at dawn (y). One might suppose that that no individual would ever prefer y to x. However, one might suppose that if the individual has an incurable disease, not fatal, but one that causes unbearable pain every day of his life, then the circumstances would make the individual end up preferring y to x.[63]

Van Hees and Wissenburg (1999) point out the difficulty of putting a reasonable person (j) in an agent's position (say $i's$), since this would require changing the reasonable person into i . They argue that "a meaningful comparison between the preferences of distinct persons in the same situation" is unlikely to be imagined (van Hees and Wissenburg 1999, 77). A further problem with approaches based on reasonable preferences is that in many situations people do not have stable or coherent preferences as traditionally assumed by rational choice theory. Sugden (2010) suggests that "a way of accounting for the value of opportunities to act on unstable and incoherent preferences" is needed (Sugden 2010, 49).

We have presented three different ways in which preferences are considering in the measurement of the opportunity aspect of freedom. We realized the importance of Sen's contribution, whose criticism against the seminal work of Pattanaik and Xu (1990) is based on the absence of preferences in their cardinality rule. Sen's work (1990, 1991, 1993b) not only establishes the importance of considering the interrelation between preferences and freedom, but also suggests how to overcome the paradox left by Pattanaik and Xu (1990). However, one of the weaknesses of Sen's approach is that it only considers actual preferences. As a result, authors such as Arrow (1995), Foster (1993), Kreps (1979) and Puppe (1996), among others, claim the importance of considering the uncertainty of individuals with respect to future preferences, in the measurement of the opportunity aspect of freedom. Finally, Pattanaik and Xu (1998), following Jones and Sugden (1982), question the fact that both the actual preference

[63] See also Arneson (1998) who analyzes this problem taking an example proposed by Jones and Sugden (1982).

approach and the uncertainty preference approach do not consider the importance of the intrinsic value of freedom of choice.

We have also noted through the analysis of the different preference-based approaches that in general they are rooted in the notion of positive freedom. Sen (1988) notes that "in traditional economic theory, insofar as freedom comes in at all, it is the negative perspective of freedom that tends to be the dominant one." However, he considers that positive freedom is no less coherent than negative freedom and he states further that the former "corresponds closely to a person being actually free to choose" (Sen 1988, 273). Bavetta (2004) argues that Sen's emphasis on positive freedom is an important element in the development of Sen's preference-based approach. Bavetta (2004) adds that economists who have contributed to the analysis of the freedom of choice that an individual has over an opportunity set "have followed suit taking Sen's thrust as the philosophical basis for the rankings. Acritically embracing the normative content of positive liberty […]" (Bavetta 2004, 32).

Although Kreps (1979) and Puppe (1996) do not explicitly mention the philosophical basis of their proposition about preferences for flexibility, this is not the case with Arrow who affirms that freedom as flexibility "crystallizes the idea that the underlying concept of freedom is freedom to choose preferences, the familiar philosophical idea of autonomy. Freedom as an ordering over sets I take to be derivative from the underlying principle of freedom as autonomy" (Arrow 1995, 11). Regarding the reasonable preferences approach, Pattanaik and Xu (1998) consider that the act of choice is based on a conscious decision, so that the agent is considered autonomous with respect to his preferences. This leads us to think that contrary to what Bavetta (2004) said, in general, the economists who proposed alternative approaches to the one proposed by Sen, questioned themselves on the philosophical basis of their proposals. The fact that the preferences for flexibility and the reasonable preferences approaches are also rooted in positive freedom does not imply that economists who have proposed them have taken the philosophical basis of these approaches for granted.

In the next section we will study the origins of preference-based approaches with the aim of establishing the tools they use from social choice theory to develop the conception of freedom in economics. We first study the plurality of interpretations of individual preferences. We show how the importance of individual preferences in social choice theory influences preference-based approaches. We then focus on analyzing how the important place given to consequences in social choice theory leads to study two aspects when assessing freedom: the process aspect

of freedom and the opportunity aspect of freedom (Sen 1990; 1993b; 2002). Finally, we examine how notions such an incompleteness and maximality that emanate from social choice theory are important in the analysis of multiple preferences and intersection rankings.

2.4 Revisiting the Origins of Preference-based Approaches

2.4.1 Versatility of Individual Preferences

> Preferences lie at the core of mainstream economics because of their connection to welfare and to choice. On the assumption that people are rational, self-interested, and knowledgeable, their preferences will both explain their choices and reflect what will benefit them. Understanding preferences better helps one understand the achievements and the challenges of economics (Hausman 2012, 8–9).

This quote merely confirms that relying on preferences to try to solve the problem of the cardinality rule proposed by Pattanaik and Xu (1990) is undoubtedly a path that should be explored by economists. As we have seen in the sub-section 2.3.3, there are different ways in which preferences can be included in the measurement of freedom of choice. But there are also different conceptions of preferences, which can be considered either as a weakness or as a strength of the approaches that consider an association between freedom of choice and preferences.

The association between freedom and the individual's preferences is related to the importance of preferences in social choice theory, where preferences are at the basis of social decisions. In essence, the importance of preferences in social choice theory is clear in the second chapter of Arrow's book: "We will therefore assume throughout this book that the behavior of an individual in making choices is describable by means of a preference scale without any cardinal significance, either individual or interpersonal" (Arrow 1951 [1963], 11). On the other hand, those preferences reflect the values of individuals rather than their tastes. "The individual orderings which enter as arguments into the social welfare function as defined here refer to the values of individuals rather than to their tastes" (Ibid, 23).

Sen (1993b) argues that the interpretation of individual preference proposed by Arrow allows us to consider that individual preference reflects both values and choices of the people. In this way, it is important to note that Sen uses Arrow's conception of individual preference ordering to justify his preference-based approach to evaluate freedom.

> The evaluation of the freedom I enjoy from a certain menu of achievements must depend to a crucial extent on how I value the elements included in that menu. The "size" of a set, or the "extent" of freedom enjoyed by a person, cannot, except in very special cases, be judged without reference to the person's values and preferences (Sen 1993b, 528).

Furthermore, the plurality of interpretations of individual preferences could call into question their role as the basis of social choice. But Sen affirms that this versatility of individuals' preferences is rather a strength:

> This plurality is a source of strength of the broad class of preference-based approaches to social choice. In different types of evaluative arguments about appropriate social decisions, diverse aspects of the individual's will, and agency are—explicitly or implicitly—considered, and the richness of the variety of interpretations permits the theory to invoke different features of the individual, depending on the context (Sen 1997a, 18).

This plurality of interpretations of individual preferences will be useful in the analysis of freedom of choice. Sen is one of the great supporters of the preference-dependent view of freedom of choice. The analysis of the debate that took place between Amartya Sen and Ian Carter is useful to understand Sen's position regarding the versatility of individual preferences.[64]

Some definitions are needed as a basis for the debate. On the one hand, Sen (1988) gives importance to the intrinsic value of freedom, the value freedom has over and above the value of the things it allows us to achieve. He insists that even if freedom also has an instrumental value (the value of freedom as means to other ends), the intrinsic value of freedom has to be considered as well. On the other hand, Carter (1995) considers that freedom has an independent value, independent of the value a person attaches to the particular things it leaves her free to do. But this independent value does not necessarily mean an intrinsic value. Carter argues that "the independent value of freedom might be intrinsic or instrumental […] The difference between freedom being instrumentally valuable and it being intrinsically valuable lies, rather, in the ultimate reducibility or irreducibility of freedom to some other good" (Carter 1995, 825–26).In 1999 Carter considers that the intrinsic value of freedom is an independent value in nature. However, this does not imply that the instrumental value of freedom must always be an

[64] This debate started thanks to a conference in honor of Amartya Sen in 1996 organized by the Center for research and instruction in politics and ethics (*Politeia*) in Milan.

independent value in nature. The instrumental value of freedom can be either independent or supervenient.[65]

Given these definitions, we present the main arguments of the debate. First, Carter affirms that the intrinsic value of freedom defended by Sen is very demanding and thus proposes a weak version of this value. This allows him to say that the independent value of freedom is exemplified not only by Sen's intrinsic value of freedom but also by the instrumental value of freedom. Carter argues that the independent value attributed to freedom is explained by the fact that individuals ignore the way in which freedom will be used to achieve outcomes that suit them.

> If freedom is instrumentally yet independently valuable, then here, just as much as where freedom is seen as intrinsically valuable, it is important to be able to assess the degrees of freedom made available by different possible opportunity sets. In short, freedom does not need to be intrinsically valuable (it need only be independently valuable) in order for degrees of freedom to be a part of the informational basis of evaluative judgments (Carter 1996, 11).

Carter does not deny at any time the intrinsic value of freedom. What he criticizes is the importance that Sen gives to this value when studying freedom. Once Carter has presented the importance of the independent value of freedom, he criticizes the preference-based approach defended by Sen. Carter takes the example that we have already mentioned of how to move around the city, in order to explain part of his criticism. We have a first option set (a stiff tricycle; hopping on one leg, rolling on dust) and a second option set (an efficient bicycle; a smart car; walking normally on two legs). Carter argues that the second option set does not offer more freedom because we prefer it. On the contrary, we prefer the second option set (in part) because it offers us more freedom.

Carter focuses his criticism against the preference-based approach on the independent value of freedom. He affirms that the preference-based approach denies the view that freedom is independently valuable because what Sen does in his example is to assess the value of particular alternatives in each option set. That is to say, Sen assesses the value of the different things freedom allows one to do. As a result, Carter concludes that the preference-based approach

[65] Carter (1999) prefers the term "non-specific value" to the term "independent value", which he used in Carter (1995, 1996). One of the reasons for this change is "to avoid the confusion of 'independent value' with 'intrinsic value'' (Carter 1999, 33). Despite this clarification, we decide to use the term "independent value" in this chapter because it is this term that Carter uses in his debate with Sen in 1996. Carter (1995) distinguishes four distinct kinds of the independent value of freedom: unconditional value, intrinsic value, non-specific instrumental value, and constitutive value.

"fails to capture the value freedom has independently of the value of those options, and, a fortiori, the value it has 'over and above' the value of those options" (Carter 1996, 12).

Consequently, Carter proposes that the basis for assessing degrees of freedom is the extreme quantities of physical action that the options provide—e.g., an efficient bicycle provides more physical action than a stiff tricycle.[66] What Carter tries to do is to demonstrate that "if freedom is independently valuable, then in terms of reason giving, freedom should come before preferences, not preferences before freedom" (Carter 1996, 12).

In reply to Carter's criticisms, Sen (1996) proposes three reasons to justify his preference-based approach: (1) An individual's preference over potential outcomes is not exclusively applied to "the things" an individual chooses, on the contrary, it is applied to the options an individual has but does not use. (2) The valuation of freedom based on preference can go beyond the valuation of "those things" that an individual gets. So individuals can have preferences over the process of choice. (3) Individuals can have preferences over alternative preferences. In other words, it is possible to consider rankings of preference rankings (meta-rankings) to express an individual's moral judgments.

In addition, Sen argues that there is no incompatibility between the preference-based approach and the intrinsic value of freedom because the value of freedom can change without a change of the things it allows one to do. "The value of freedom can change when there is a significantly improved unchosen option, or a preferred process, or a better opportunity in terms of a different preference the person could have preferred to have" (Sen 1996, 110). Sen's reply would have been incomplete if he did not take up the independent value of freedom proposed by Carter. He expresses his disagreement with the idea that freedom is independently valuable because this could mean missing the opportunity to assess the value of freedom according to the alternatives over which we can freely choose. It is important to note that Sen considers that if we assumed the premise that freedom is independently valuable, it may lead us to assess freedom according to the number of available alternatives in an option set.

Therefore, Sen defends the relation between the value of freedom and the value of the alternatives over which individuals can choose. To better understand this relation, we present another of Sen's examples in a slightly modified form: a woman decides to spend her day off

[66] See Carter (1999), in order to study the empirical approach to measuring freedom. According to this approach the extent of someone's freedom is a function of the extent of an action available to her, in "sheer quantitative terms." (Ibid, 170). This approach "rests on the view that purely physical descriptions of actions are the relevant kinds of descriptions when it comes to measuring degrees of overall freedom" (Ibid, 212).

at home after a hard week in her work; she could have chosen to go out for a walk, but she does not do so. Now consider two counterfactual scenarios:

1. A bully comes and tells her, "You must not go out; you must stay at home."
2. A bully comes and tells her, "Take a dip in the sewers" (something that she would have avoided choosing at all costs).

Even if in the first scenario, this woman ends up doing what she had planned to do (stay at home), she lost the option of doing something else. This allows Sen to say that freedom involves more than the value of the different things it allows us to do, which reinforces the importance of the intrinsic value of freedom that he defends. Another point that is important to highlight from this example is the fact that in the second scenario the violation of freedom is more important that in the first case. In essence, in the second scenario this woman is forced to do something that she would never have chosen to do on her day off (take a dip in the sewers). Therefore, she has obviously less freedom to do what she prefers to do on her day off in this second scenario than in the first scenario, in which she ends up doing what she had in mind, even if she did not freely choose it. The comparison of these two scenarios allows Sen to show that preference is pertinent in assessing the value of freedom we have, and this assessment is influenced by the things we achieve.

In brief, Sen seeks to show that Carter makes a mistake when he affirms that freedom is independently valuable. So, in addition to the example explained above, he proposes a distinction between two forms of dependence of the value of freedom in order to complete his argument against Carter: (1) the value of freedom depends, inter alia, on the value of the objects achieved, and (2) the value of freedom depends exclusively on the value of the objects achieved. (Sen 1996, 111). Sen argues that Carter in his effort to reject the exclusive dependence of the value of freedom on the value of the objects achieved, also ends up rejecting the inter alia dependence of the value of freedom on the value of the objects achieved.

The irreconcilable positions between Sen and Carter come from the way they both define the value of freedom. On the one hand, Sen relies on the plurality of individual preferences and in different examples, such as the woman who decides to take a day off, show that the value of no-chosen alternatives is important to the evaluation of a person's freedom. On the other hand, Carter considers that Sen's preference-based approach fails to capture what he calls the independent value of freedom, which means that Sen "reduces the value of having a certain measure of freedom to the values of the specific things that one is free to do […]" (Carter 1999,

127). Carter rejects all the value-based approaches for the reasons stated above. He proposes that in order to assess degrees of freedom what is important is the "physical extension" of the actions which the alternatives allow.

This debate is still topical in studies on the assessment of degrees of freedom, such as in Binder (2019). She demonstrates in opposition to Carter that value-based approaches can capture the independent value of freedom.[67] She takes another example proposed by Sen in his reply to Carter, the example of a person that is fasting or starving depending on the presence of the option of eating. Sen considers that having the possibility to eat will give the person the "freedom to reject." This means that this person can choose to fast because she has the option of eating, which she can deliberately reject. Binder argues that a value-based approach can capture the independent value of freedom, since the importance that a person gives to the rejected option depends on her values. However, this does not necessarily imply that the independent value of freedom "can be captured by preference-based approaches, which are a special class of value-based approaches" (Binder 2019, 123).

This means that the existing preference-based approaches cannot explain certain instances of the independent value of freedom, such as the "freedom to reject."[68] Binder proposes a solution. In essence, it is a refined approach to capture the interdependencies of alternatives in a person's opportunity set. "Preferences over refined alternatives specify both the alternative which is chosen and the set from which it is chosen" (Binder 2019, 124). In brief, this solution confirms that a preference-based approach can capture different instances of the independent value of freedom under a given framework, and it is useful to demonstrate that value-based approaches do not fail to take the independent value of freedom into account.[69]

[67] Depending on how the value of alternatives freedoms is identified, we can distinguish between a subjective and an objective line of value-based approaches. The subjective line identifies the value of an alternative based on an individual's preferences. On the other hand, the objective line identifies the value of particular freedoms based on some standard, which is independent of the individual concerned (Binder 2019, 25–6).

[68] Binder (2019) studies one of the instances of the independent value of freedom. It is about freedom's agency value: "the value freedom has in virtue of allowing a person to reflect upon the relative priority of her various identity parts in the moment of choice (represented by her all-things-considered preference)" (Binder 2019, 94–5).

[69] The preferences over refined alternatives proposed by Binder would allow "one to distinguish between starving or fasting, depending on whether or not the option to eat has been available (and deliberatively rejected)" (Binder 2019, 124).

2.4.2 Two Aspects of Freedom and Its Connection to Preferences

In social choice theory preferences play an important role. But a central place also given to consequences is highlighted by Sen

> The philosophical approach underlying social choice theory has strong consequentialist sympathies. Individual preferences are defined over the space of social states, the exercise of aggregation concentrates on social states, and decisions about policies, institutions, rules and so on are seen in the terms of appraisal of the states they generate (Sen 1997a, 24).

The important place given to consequences in social choice theory leads us to consider the process aspect of freedom concerned with the procedure of autonomous choice, and the opportunity aspect of freedom concerned with the real opportunities that individuals have in order to achieve things that they value (Sen 1993b, 2002). The process aspect includes two features, (i) autonomy of decision or self-decision when making a choice, and (ii) immunity from encroaching activities that means the absence of interference by others. The immunity is related to the concept of negative freedom. Sen points out that Berlin's definition of negative freedom "takes note of the various parts that others play in making a person unable to do something, and therefore goes well beyond the "'immunity' component of the process" (Sen 1993b, 524). Sen uses a narrow notion of negative freedom. The opportunity aspect considers "the opportunity of achieving the best that can be achieved, but may extend that concern by taking some supplementary note of the range of opportunities offered" (Ibid, 525).

We have already discussed (sub-sections 2.3.3 and 2.4.1) the relationship between the opportunity aspect of freedom and the individual's preferences. We now focus on the process aspect of freedom and its connection with the opportunity aspect. The process aspect of freedom also has a connection with preferences. Sen sees two ways in which preferences are pertinent in judging processes: "(1) Personal process concern: individuals may have preferences over the processes that occur in their own lives; (2) Systemic process concern: they may also have preferences over the processes that operate as general rules in the working of the society" (Sen 2002, 624). Personal process concern describes how individuals' preferences may be important to make their choices, lead their lives, and so on. Systematic process concern describes how individuals' preferences may also be important for their conceptions of social institutions and rules of social behavior. When evaluating the essential attributes and the magnitude of personal

freedom, personal process concern must be considered. Similarly, systemic process concern may be useful in assessing the state of a process-oriented freedom, which is "judged in terms of the fulfillment or violation of systemic rules relating to processes" (Ibid, 625).

Sen distinguishes the opportunity aspect and the process aspect of freedom, but he does not rule out that these two aspects have a connection.

> The value of [an opportunity] set need not invariably be identified with the value of the best – or the chosen – element of it. Importance can also be attached to having opportunities that are *not* taken up. This is a natural direction to go if the *process* through which outcomes are generated is of a significance of its own. (Sen 1997c, 202).

Consider a well-known example in the literature to show the connection between the importance of preference over no-chosen options and that over processes (Sen 1988; 1993a; 1996): (1) a person who is starving because of poverty, and (2) a person who is fasting. In both cases the person does not eat. But they differ with respect to the process of choice. The person who is starving does not have the option of eating. The person who voluntarily chooses the fasting option, does it for political or religious reasons. Another difference refers to the options available in each case. In the first case the option to eat is not available while in the second case the option to eat is open for choice. If the option to eat is removed, in the second then the person's preference would be affected, since her preference was for "fasting," (with eating as open for choice but rejected) not for "starving."[70] It is worth emphasizing that the opportunity aspect when referring to the freedom that a person has or achieves would be said to be rooted in positive freedom as defined by Berlin. The process aspect takes into account both positive freedom, through the autonomy component, and negative freedom, through the immunity component. It seems that these two aspects of freedom are part of a broader definition. Sen (1988; 2002) proposes a broader definition of positive freedom: "the person's ability to do the things in question *taking everything into account* (including external restraints as well as internal limitations)." (Sen 2002, 586).

[70] Sen (1993a) uses this example to show one of the ways in which the contraction consistency condition (α) or "the Chernoff condition" is violated. He calls this case "the freedom to reject." The condition α is defined in rational choice theory as follows: "If some element of subset S_1 of S_2 is best in S_2, then it is best in S_1" (Sen 1970 [2017], 63).

2.4.3 Multiple Preferences Approaches and Intersection Ranking

> One reason why incompleteness tends to arise in individual valuation is the fact that a person often has reasons to entertain the idea of having different preferences. Such plurality of valuations can be treated either as a person's having "multiple preferences," or – in an informationally reduced form – as her having an incomplete ranking that reflects the intersection of different preferences (Sen 2002, 599).

In assessing freedom, the two conceptions of incompleteness and maximality cannot be overlooked. It is important to note that these conceptions emanate from social choice theory. We first define the two terms in order to understand their role in social choice theory, and then study the way both terms can be important in the analysis of multiple preferences.

In order to define incompleteness, we have to consider one of the ingredients of collective choice, namely ordering. According to Arrow "each individual in the community has a definite 'ordering' of all conceivable social states, in terms of their desirability to him […] it is simply assumed that the individual orders all social states by whatever standards he deems relevant" (Arrow 1951 [1963], 17). To be considered as an ordering, three characteristics must be satisfied: transitivity, reflexibility and completeness.

Let $x\ R\ y$ represent a binary relation of 'weak preference' ('x is at least as good as y'). To specify this binary relation over a set S is to identify a subset R of the square $S \times S$, defined as the set of all ordered pairs (x, y) so that x and y both belong to S.

1. Transitivity: $\forall x, y, z \in S: (x\ R\ y\ \&\ y\ R\ z) \rightarrow x\ R\ z$
2. Reflexivity: $\forall x \in S: x\ R\ x$
3. Completeness: $\forall x, y \in S: (x \neq y) \rightarrow (x\ R\ y \veebar y\ R\ x)$ (Sen 1970 [2017, 52–3).]

We focus on the analysis of the third characteristic. According to Sen, "a man with a preference relation that is complete knows his mind in choices over every pair" (Ibid, 47). There is a distinction between "indifference" and "lack of completeness." Sen (1970 [2017]) explains it by two statements:

(1) x is at least as good as y

(2) y is at least as good as x

When we talk about "indifference" both statements are true, whereas when we consider "lack of completeness" neither is true. The preference ranking relation that satisfies both transitivity and reflexivity but does not necessarily satisfy completeness is called a quasi-ordering.[71]

Incompleteness in a preference ranking relation can result not only from "assertive" incompleteness, but also from "tentative" incompleteness. The former is "when some pair of alternatives is asserted to be 'non-rankable'." The latter is "when some pairs of alternatives are not yet ranked (though they may all get ranked with more deliberation or information)" (Sen 1997b, 182).

The conception of incompleteness in a preference ranking cannot be studied in isolation from the distinction between "optimality" and "maximality". We present this distinction formally.

A choice function based on optimization can be expressed as:

$$B(S,R) = [x|x \in S \,\&\, for\ all\ \gamma \in S: xR\gamma]$$

From this function, we can choose a "best" element "from each menu set S, according to a weak preference relation R (interpreted as 'preferred or indifferent to'), which ranks the set of available alternatives X of which each 'menu' S is a non-empty subset." (Sen 1997b, 172).

With respect to a function based on maximization, the formal expression is as follows:

$$M(S,R) = [x|x \in S \,\&\, for\ no\ \gamma \in S: \gamma Px]$$

This means "for an element of S to qualify for the maximal set $M(S,R)$, no other alternative in S must be strictly preferred to it." (Sen 1997b, 182).

The basic difference between maximality and optimality comes from the possibility of an incomplete preference ranking. In order to better understand this distinction, we present the well-known example of Buridan's donkey. This is a hungry donkey who stands between two equidistant and attractive haystacks. The donkey starves to death from dithering in the choice of the two haystacks. There is this tragic ending because it is "seeking optimality when it could not figure out which of the two haystacks facing it was the better one to go for." (Sen 2017, 454). It could not figure out if one of the haystacks is, for example, more appetizing than the other. It would have been better to go to either one of the haystacks rather than die from starvation. We are facing here a case of incompleteness of the ordering of the two haystacks, and incapacity to identify the best alternative (lack of optimality). The two haystacks must be

[71] See Chapter 1* Preference Relation in Sen (1970 [2017]) for the formal properties of a quasi-ordering.

considered as maximal, and the choice of either one would have been a better option than to die from hunger.[72]

Once the definitions of incompleteness and maximality have been presented, we analyze their relationship with the value of freedom. We have to consider the opportunity aspect of freedom, which refers to the opportunity of accomplishing the best that can be accomplished.[73] Given that a person's preference may be incomplete, her choice cannot be seen as the "best" for her, "even if she sticks to maximizing behavior, which would only require that the chosen alternative be 'maximal.'" (Sen 2002, 608). It means that the chosen alternative is prevented from being worse than the remaining options.

Sen argues that an incomplete preference ranking is not a limitation in evaluating opportunity "because there are often compelling grounds for not being able to rank every alternative vis-à-vis every other, and also because it is possible to make effective use of incomplete orderings in making rational choices" (Sen 2002, 609). The Buridan's donkey example shows us what happens when an incomplete ordering is considered as a decisional impasse. For Sen it is important to recognize that incompleteness is "an actual—and may even be a common—outcome of reasoned analysis of ethical and political evaluation." (Sen 2017, 458).

It is important to note that the evaluation of freedom does not lie in being able to rank each set of options vis-à-vis the others, but rather in making use of as much reasoning as possible to rank different alternatives in a very definite way. For instance, a complete ordering of freedom is not necessary in order to recognize the unfreedom that extreme poverty can cause in a person's life.

When Sen (2002) studies the multiple preferences that an individual can have, he establishes a relationship with the conception of incompleteness. The multiple preferences of an individual may bring up incompleteness. "A person can have reason to consider preference rankings that she does not actually have and can sometimes even have reason to prefer having a preference ranking other than the one she has." (Sen 2002, 669). This latter quote leads us to the conception of meta-rankings, the ranking of preference rankings.

> Let X be the set of alternative and mutually exclusive combinations of actions under consideration, and let Y be the set of rankings of the elements of X. A ranking of

[72] In another interpretation of Buridan's donkey it is assumed that the donkey is indifferent between the two haystacks, so it could not have any reason to choose one of the haystacks.

[73] A person's preferences may be incomplete, and this can affect the real opportunities that she has to achieve her objectives. Only the opportunity aspect of freedom and its relation to incompleteness are considered in our study.

> the set Y (consisting of action rankings) will be called a meta-ranking of action-set X (Sen 1977a, 337).

Sen (2002) argues that the notion of meta-rankings is related to a person's autonomy, and he sees three ways in which a person's autonomy plays a relevant role in evaluating opportunities. First, a person must be able to decide what importance she gives to the preference that she prefers to have, instead of any other preference she would have preferred to have. Second, a person must have the freedom to re-examine the preferences she has. Third, a person may also resent if others take her preferences as "given." In short, autonomy is focused on what a person can do and also on what others must not take as given.

Sen distinguishes three cases in which meta-rankings are closely related to a person's autonomy. First, "a person's meta-ranking may be complete enough to yield the judgment that a particular preference ordering is the 'best' one to have. And yet the person may not manage to 'lick' her preference into shape […]" (Sen 2002, 670). It means that if a person has a particular preference ordering, it does not imply that she cannot decide on the importance given to that preference, rather than to another. Second, "for a counterfactual preference R_j' to be "relevant" to a person's assessment of options, that preference ranking R_j' need not necessarily be meta-preferred to the actual ranking R_j." (Sen 2002, 670). It is important to see that incompleteness can also be present in a meta-ranking. We can also invoke the distinction between maximality and optimality when we study meta-ranking. Third, a person might "be impressed by the relevance of the other preference orderings, and in particular might refuse to accept that in the evaluation of her freedom, the other preferences could be simply ignored." (Sen 2002, 671). In essence, it is the person and not others in her place that decides the preference ordering that she wants to hold and use, and what alternative action to choose.

In brief, Sen tries to show that in order to assess a person's opportunities and freedom, it is important to take into account not only a person's actual preference but also a person's relevant counterfactual preferences. In this way, the person's autonomy will be respected and also the possibility that a person could have chosen to have another preference. Frankfurt in his study about the structure of person's will, had proposed a sort of meta-ranking. He suggests that what makes a person different from other creatures is not only to have what he calls "first-order desires", i.e., "simply desires to do or not to do one thing or another," but also to have "the capacity for reflective self-evaluation that is manifested in the formation of the second-order

desires." Someone has a second-order desire "when he wants to have or not to have a certain desire of the first order." (Frankfurt 1971, 7).[74]

Beyond the possibility that multiple preference orderings can be meta-ranked vis-à-vis each other, they influence the study of opportunity and the relevance of alternatives. One way to study how multiple preferences can deal with the analysis of opportunity and the importance of alternatives is "to take the partial ordering that emerges from the 'intersection' of different preferences (or valuational functions) that a person may regard as relevant for the assessment of her freedom." (Sen 2002, 671).[75] This intersection of quasi-orderings let us consider once again the incompleteness of the ranking of opportunity freedom. The incomplete ranking may be adequate to capture the presence and relevance of multiple preferences.

Sen's Intersection approach first takes "the intersection R^I of multiple preferences, which would in general be a partial quasi-ordering, and then apply Elementary Option Superiority (EOS) or Elementary Correspondence Superiority (ECS)…"(Sen 2002, 612–13).[76] It means that if we compare two opportunity sets A and B, we can examine whether there is some alternative in set A such that it is at least as good as every alternative in set B. If this is the case, we are in presence of the condition called Elementary Option Superiority.

Formally: Elementary Option Superiority (EOS): If for some x in A, we have $xR_j\gamma$ for all γ in B, then $A\ R^*\ B$.

This condition states that to say that A has at least as much elementary value, it is sufficient to demonstrate that some option in A is at least as good as any option in B (Sen 2002, 667).

This less demanding condition is called Elementary Correspondence Superiority and examines whether for every alternative in set B there is superior alternative in A. So, it is not necessary to be the same alternative that is considered to be superior in each case, like in the case of EOS.

Formally: Elementary Correspondence Superiority (ECS): If there is a functional correspondence $k(.)$ from B to A such that for every element γ in B, $k(\gamma)\ R_j\ \gamma$, then $A\ R^*\ B$.

[74] Frankfurt is not limited to second-order desires and is aware of the complexity of people. A person may have higher orders than the second-order (Frankfurt 1971, 16).

[75] A partial ordering satisfies reflexivity, transitivity, and anti-symmetry. "Anti-symmetry: $\forall x, y \in S: (x\ R\ y\ \&\ y\ R\ x) \rightarrow x = y$" (Sen 1970 [2017], 53).

[76] There are other intersection approaches, which do not reach the same conclusions as Sen. See Foster (1993) and Binder (2019).

This condition states that it is sufficient to show that every alternative in B is matched by some alternative or other in A (Sen 2002, 668).

It is important to note that both the EOS and ECS conditions "will, in general, yield an incomplete ranking of opportunity, given the incompleteness of the underlying preference quasi-ordering." (Sen 2002, 611)[77]

An example given by Sen can be useful in order to understand his intersection quasi-ordering approach:

> A person with musical talent but who is not averse to affluence considers three options: becoming a full-time musician, which in that society will also yield a very low income (x); doing a bit of music as a sideline with a primary job in another field that yields a moderately high salary (γ); and becoming a full-time businessman who has no time for music but much affluence otherwise (z) (Sen 2002, 614).

Sen assumes two alternative rankings in relationship with the musical orientation and the affluential orientation of the person. The musical orientation of the person places the options in a decreasing way: x, γ, z. The affluential orientation of the person places the options in a decreasing way: z, γ, x. It is important to note the incompleteness of the intersection ranking, i.e., the set of preference orderings can rank the three alternatives in a quite definite way.

We can depict what has been said above as follows:

Table 3. Incompleteness of the intersection ranking

Musical orientation	Affluential orientation
x	z
γ	γ
z	x

→

$xN\gamma$
xNz
γNz

If R is incomplete with regard to some $x, \gamma \in X$, we write $xN\gamma$. For all $x, \gamma \in X$ if $xN\gamma$ then γNx.[78]

[77] A comparison between EOS and ECS is made in Sen (2002), particularly in Chapter 22.
[78] This terminology is based on the interpretation by Binder (2019) of Sen's example.

Sen compares the two opportunity sets $A = \{x, z\}$ and $B = \{x, \gamma, z\}$. He aims to elucidate if set A offers as much agency freedom as set B. "A person's 'agency freedom' refers to what the person is free to do and achieve in pursuit of whatever goals or values she regards as important." (Sen 1985, 203). We can observe that if γ is removed from set A, then Sen (2002) argues that set A offers less agency freedom than set B. In essence, a person with musical orientation could have chosen γ over z, and a person with affluential orientation could have chosen γ over x. Therefore, in both cases, by having the option of choosing γ, a person prevents the possible worst. According to Sen's intersection approach, the fact of removing an option from one of the sets affects a person's agency freedom.[79]

2.5 Conclusions

We make a few general remarks about the most important issues covered in this chapter.

1. Freedom as a triadic relation is at the basis of the difference between freedom and freedom of choice. Carter (2004) suggests that both the preference-based approach of freedom and the non-preference-based approach of freedom fail to propose a solution for the dilemma presented by Pattanaik and Xu (1990).
2. The preference-based approaches are rooted in the notion of positive freedom. We consider that in general the economists who have contributed to the analysis of the freedom of choice that an individual has over an opportunity set do not take the philosophical basis of their approaches for granted. This leads them to propose three measures of freedom of choice: freedom in terms of a person's actual preferences, freedom in terms of future preferences and freedom in terms of reasonable preferences.
3. The connection between freedom and an individual's preferences is a consequence of the importance of preferences in social choice theory, where they are at the basis of social decisions. It is important to say that Sen defends the plurality of the concept of preferences.
4. The debate between Carter and Sen shows us the importance of the difference between the intrinsic and instrumental value of freedom. The subject of the debate is to determine if the preference-based approach is appropriate to study individual freedom.

[79] It is important to say that if we analyze this example through Foster's (1993) intersection approach, we do not obtain the same results.

5. The important place that consequences have in social choice theory leads us to consider two aspects of freedom: the process aspect and the opportunity aspect.
6. Incompleteness in a preference ranking has to be studied together with the distinction between optimality and maximality. These conceptions come from social choice theory, and they play an important role in the preference rankings for the evaluation of freedom.
7. The partial orderings that emerge from the intersection of multiple preferences constitute one way to assessing an individual's freedom. The plurality of preferences is at the basis of incompleteness.

In this chapter, we have attempted to explore the tools that social choice theory brings to the study of preference-based approaches to measuring freedom of choice. In the next chapter, we focus on studying the role of value judgments in the construction and interpretation of Arrow's impossibility theorem. This will allow us, among other things, to continue studying the importance that Arrow gives to individual preferences in social decisions.

Chapter 3

Revisiting the Role of Value Judgments in Arrow's Impossibility Theorem

---

Abstract

This chapter provides an analysis of the role of value judgments in the elaboration and interpretation of Arrow's impossibility theorem, that played an important role in the progress of normative economics. We first consider the "value-relevance" of Arrow's impossibility theorem. In other words, we identify the evaluative judgments that Arrow makes regarding the methods and approaches he chooses to employ in order to obtain his impossibility theorem. Then we analyze Arrow's impossibility theorem under the prism of Mongin's work (1999; 2002; 2006b). We consider three reasons why Mongin's contributions shed light on Arrow's position vis-à-vis the role of value judgments in his impossibility theorem. This allows us to analyze the value judgments that Arrow makes himself, to analyze not only each one of the conditions of his theorem but also when relaxing one of the conditions to solve the impossibility result.

3.1 Introduction

Arrow's impossibility theorem is without any doubt a seminal contribution to the development of modern social choice theory and to the progress of normative economics. Since its appearance in 1951, this impossibility theorem has generated an enormous interest because it shows that a group of intuitive conditions cannot be simultaneously satisfied by any social choice procedure. The most striking result of the impossibility theorem is that only dictatorship would prevent inconsistencies. This implies two important consequences: "(1) in politics, an extreme sacrifice of participatory decisions, and (2) in welfare economics, a gross inability to be sensitive to the heterogeneous interests of a diverse population." (Sen 1999, 351).

The aim of this chapter is to analyze the role of value judgment in Arrow's elaboration and interpretation of his impossibility theorem. This fits well into one of the objectives of this dissertation, which is to know to what extent the economist can make value judgments in normative economics theories. We start with inquiring how Arrow's own values and the cultural context influenced his methodological choice. This allows us to identify what Weber calls the "value-relevance" of Arrow's impossibility theorem. However, Arrow's value judgments play an important role not only when he decides how to conduct of inquiry but also in the formulation of the conditions of his impossibility theorem.

In order to study the "value-relevance" of Arrow's impossibility theorem, we focus on interviews made with Arrow by his colleagues and on scientific articles where Arrow explains the origins of his impossibility theorem. This allows us to examine the evolution of Arrow's thinking. From his interest in the different expectations that owners have in a firm to his interest to formalize the idea that voters have preferences for certain parties. These ideas lead him to study the meaning of preferences for a collectivity of individuals until reaching an impossibility result (Arrow 2014b).

We analyze Arrow's impossibility theorem under the prism of Mongin's work (1999; 2002; 2006b). We consider three reasons why Mongin's approach sheds light on the Arrow's position

vis-à-vis the role of value judgments in his impossibility theorem. First, Mongin (1999) analyzes the extent to which the economist can avoid making value judgments in normative economics. Second, Mongin (2002) examines the importance of Arrow's impossibility theorem in the progress of normative economics. Third, Mongin (2006b) proposes a value neutrality classification (presented in Chapter 1) that allow us to clarify the role of value judgments in economics.

We take the interpretations proposed by Mongin in 1999 in order to study how Arrow considers value judgments in his impossibility theorem as the starting point. These interpretations cannot be studied in isolation. For this reason, we consider Mongin's analysis in 2002 of the impact of Arrow's development of social choice theory in the progress of normative economics. We also examine Mongin's 2006b hypothesis that Arrow can be considered as a weak neutrality theorist, which means that Arrow considers that it is logically possible for an economist *qua economist* to make value judgments, and that there are occasions in which he might, and even should, make value judgments.

This analysis of the value judgments in Arrow's elaboration and interpretation of the impossibility theorem leads us to examine how to escape from this result. We do not intend to make an exhaustive study of the different propositions made by economists to relax some of the conditions proposed by Arrow. We focus on studying the value judgments that lead Arrow to propose alternative solutions in order to escape from his impossibility result. In a nutshell, the main contribution of this chapter is to show that Mongin's value neutrality classification studied in the first chapter of this dissertation is useful to determine that Arrow cannot dispense with making value judgments when stating his impossibility theorem and its possible solutions.

The chapter is organized as follows. Section 2 presents the methodological value judgments made by Arrow in order to choose the research topic, the method of investigation and the standards of validity. Section 3 discusses Arrow's position with respect to the role of value judgments in his impossibility theorem and the value judgments behind Arrow's proposals to escape from his impossibility result. Section 4 concludes by listing the most important results of our analysis about the "value-relevance" and the value-ladenness of Arrow's impossibility theorem.

3.2 The "Value-relevance" of Arrow's Impossibility Theorem

In order to study the methodological choices made by Arrow, we examine the "value-relevance" of his impossibility theorem. "Value-relevance" is a concept developed by Weber that we have already studied in the first chapter of this dissertation. The importance of this concept lies in the fact that it implies that the researcher's evaluative ideas are essential to the "choice of the object of the investigation and the extent or depth to which this investigation attempts to penetrate into the infinite causal web." In other words, the evaluative ideas that guide the researcher, that it is to say "his personal belief" and "the refraction of values in the prism of his mind," are essential to determine the important or unimportant aspects of the phenomenon under study (Weber 1904 [1949], 82).

We consider that the definition of methodological value judgments in economics proposed by Boumans and Davis (2010) reflects what Weber means by his definition of "value-relevance." They define them as: "evaluative judgments which economists make with respect to the methods and approaches they choose to employ in their investigation of the economy." (Ibid, 171). They propose three types of choices that economists have to make in order to practice economics: (1) the choice of the research topic; (2) the choice of the method to be used in studying that research topic; and (3) the choice of the criteria, standards, and norms used to evaluate the validity of the research results (Ibid).

We aim to identify the methodological value judgments made by Arrow in order to obtain his impossibility theorem. We study the storytelling proposed by Arrow regarding the elaboration of his impossibility theorem approximately 30 years after the first publication of *Social Choice and Individual Values* (SCIV).

Among the resources that we use to achieve this objective, we have the first volume of the *Collected Papers of Arrow* (the foreword of his 1950 paper "A Difficulty in the Concept of Social Welfare"). We also consider the interview that Professor Arrow gave to Jerry Kelly (one of his former students at Harvard) on March 4, 1986. It was published in *Social Choice and Welfare* in 1987 and reproduced in chapter 13 of the *Handbook of Social Choice and Welfare* (Volume II) in 2011. In 1987, George R. Feiwel published two volumes in honor of Arrow's extensive contribution to modern economic theory and the theory of economic policy. In these volumes, there are two interviews (chapter 2 of the first volume and chapter 23 of the second

volume) that Arrow gave to Feiwel regarding the origin of his contributions and the topics that interested him at that time. In a more recent interview conducted in 2011 by Jane Hibbard as a part of the Stanford Oral History Program, Arrow gives the same storytelling about the elaboration of the impossibility theorem. Beyond the interviews, Arrow (1991) writes about the origins of his impossibility theorem in a book dedicated to the history of mathematical programming (this article is reprinted in Maskin and Sen (2014)).

3.2.1 The Choice of the Research Topic

This first choice that economists have to make produces a point of entry to their research. The methodological value judgments economists make about their point of entry may sometimes be linked to their ethical views, but they may also not be (Boumans and Davis 2010). In our case, we argue that Arrow has three main points of entry that lead him to obtain his controversial result.

The first point of entry that we identify is the interest that Arrow shows for the standard theory of the firm which is based on the hypothesis that the objective of the firm is to maximize profits. Arrow himself states that the first "serious encounter with the idea of social choice" was at the Columbia University in a discussion in 1946 with his Ph.D. supervisor, professor Albert G. Hart (Arrow 1991 [2014b], 145).[80] Hart brought Arrow's attention to the diverse expectations that shareholders have when choosing an investment project in order to maximize expected profit. What leads Arrow to question how firms make their decisions is his interest in redoing Hicks's *Value and Capital* (1939).[81]

Arrow is specifically concerned with the different expectations that owners (shareholders) might have in their interest to maximize expected profit. They would not agree on the type of investment that they would make in order to guarantee production in the future. "The natural assumption was that one investment policy would be chosen over another if a majority of the stockholders (weighted by the number of shares) preferred the first to the second." (Arrow 1991

[80] Arrow obtained a fellowship in 1941 at Columbia University to do a Ph.D. in Economics. He interrupted his research to do military service in the U.S. Army Air Corps at the Headquarters Weather Division. After the war, he returned to Columbia University in 1946.

[81] Redoing Hicks properly was the first topic that Arrow considered for his Ph.D. dissertation. However, he disregarded this thesis topic because it was a vast subject. "I wanted to combine it (*Hicks' work*) with Samuelson's (1941, 1942) stability theory […] Maybe I would add some stochastic elements to the story […] Well, it was a lifetime of work, really; it was a very unrealistic thesis" (Arrow, Sen, and Suzumura 2011, 8).

[2014b], 146). In other words, Arrow considers the majority vote to determine the preferred type of investment of shareholders, but he realizes that this relation is not transitive. In the following quote, Arrow clarifies this point:

> […] They choose according to majority voting, allowing for the number of shares. So one man may have 100 votes, another man has 20 votes, but allowing for that I realized this wouldn't work […] if they choose investment project A over investment project B, and investment project B is better than investment project C, and you want to maybe say that investment project A is better than investment project C. Well, it turns out if you use majority voting, that won't work (Hibbard 2011, 47).

This interest in the different views of shareholders when making a decision about the investment project that maximizes profit could have become the subject of Arrow's dissertation. But he abandoned the whole dissertation in 1947 because he could not believe that it was an original subject, and he had the conviction that he had heard of this paradox before.[82]

The second point of entry that we distinguish is related to electoral politics. Arrow started to explore the idea that voters have preferences over candidates (or parties). He proposes a model in which parties are put in order on a left-right scale. Then the supporters of any one party would prefer a party close to their preferred party over one further from their preferred party. It leads Arrow to think that under these conditions, majority voting would determine an ordering. But again Arrow did not pursue the matter because about a month after these reflections, he found the same idea in a paper by Duncan Black in the *Journal of Political Economy* (Arrow 1991 [2014b]; Arrow, Sen, and Suzumura 2011; Feiwel 1987a). This all happened when Arrow was at the Cowles Commission for Research in Economics at Chicago (1947–1948). Arrow considers that this kind of reflection distracted him from a serious economic work, "specifically using general equilibrium theory to develop a workable model as a basis for econometric analysis." (Arrow 1991 [2014b], 147).

The third point of entry that we consider is related to game theory. Game theory was a new topic in 1949, and many specialists invited by the Rand Corporation, among them Arrow, discussed conceptual and technical aspects of this theory (Arrow 1991 [2014b]; Arrow, Sen, and Suzumura 2011; Feiwel 1987a; Hibbard 2011). Among the invited members of the Rand, there was the philosopher Olaf Helmer. He made Arrow realize that his thoughts about voting

[82] Arrow admits that he omitted an important element in his analysis of the decision of shareholders. He did not take into account that the shareholders can sell their stock (this point is mentioned in the interviews that Arrow has with Jerry Kelly (1986) and George R. Feiwel (1987a)). In the interview with Feiwel (1987a), Arrows states "It is not quite true that you have to go to the majority […] so the model of the firm was not really right because I did not recognize stock transfer" (Feiwel 1987a, 192).

systems were not without originality, and that they were in no way less important than an econometric study on general equilibrium theory. Helmer was troubled by the application of game theory when players are nations and not individuals, and he raised his doubt to Arrow.

Arrow described Helmer's questioning as follows:

> They were taking game theory and applying it especially to Soviet−U.S. relations: diplomatic conflict, potential tactical situations, war. However, the payoff functions were defined in terms of utility functions […], and these were derived on the basis of the individual. The trouble was, the Soviet Union and the United States were not individuals. What is the meaning of this? […] (Arrow, Sen, and Suzumura 2011, 13–14).

Arrow replied to Helmer: "Economists have thought about that, and it's really explained by Bergson's social welfare function." (Arrow, Sen, and Suzumura 2011, 14). Helmer said that if that were the case, it would be a good exercise to write up an exposition of how Bergson's social welfare function settled this. This discussion between Helmer and Arrow seems to be a simple anecdote, but we consider that it allowed Arrow to have a light-bulb moment in the development of his final dissertation. At least this is what Arrow tells us in different interviews in the 1980s and his later writings in Arrow (1991 [2014b]). In a more recent interview with Hibbard (2011), Arrow reiterates the importance of Helmer's questioning in the development of his final dissertation: "I really owe it to Helmer who posed the question. I'm giving him credit for that […]" (Hibbard 2011, 58).

We can say that the first point of entry is determined by the academic context because Arrow shows his interest in creating a statistical model based on the work of John Hicks. Arrow studied statistics as a graduate student in mathematics at Columbia University before the Second World War. For this reason it is not surprising that he seeks to apply his knowledge of mathematics and statistics to understand economic theory. The other two points of entry are more associated with the geopolitical context characterized by the beginning of the Cold War.[83] In an interview with Kristen Monroe and Nicholas Lampros, Arrow states that one of the main concerns of the Rand Corporation was

> the possibility of nuclear annihilation and any other possibilities with regard to our conflict with the Soviet Union, which I felt was personally very important […] One of the questions that came up was: "Well, what do we mean by the United States'

[83] Amadae (2003) considers that the context of the Cold War played an essential role in the creation of Arrow's impossibility theorem. She argues that the impossibility theorem "plays a central role in defining the Cold War consensus on the philosophical foundations of economic and political liberalism" (Amadae 2003, 112).

> interests?" The United States is an abstraction. The country is composed of a lot of people, with different interests, just as the Soviet Union is (Arrow, Monroe, and Lampros 2017, 90).[84]

We claim that the three points of entry discussed in this section help us understand how Arrow fine-tuned his thesis topic until he got to his impossibility theorem. Each point of entry reflects the way in which the thesis topic was taking shape in Arrow's mind. We consider that Arrow himself summarizes well this process: "I also had the habit of trying to get things into logical order. That's my talent. I like to connect things which seem to belong to different realms. I see the parallels. I'm rather good at that" (Arrow, Monroe, and Lampros 2017, 90).

3.2.2 The Choice of the Method of Investigation

Arrow did not come from a scholarly family. His mother graduated from high school and his father got a degree from New York University Business College. From an early age Arrow was a good student. He skipped grades in school and finished high school at fourteen, the Townsend-Harris High School in New York City where he developed a great interest and skill in mathematics. Arrow was a voracious reader and he studied logic on his own even at high school (Arrow, Sen, and Suzumura 2011; Feiwel 1987b; Hibbard 2011).

When Arrow went to the City College of New York, he got fascinated by statistics and logic. He read original literature in statistics (Hibbard 2011). Arrow took a formal course on the calculus of relations with the great Polish logician, Alfred Tarski. This course was probably based on the article published by Tarski in 1941, *On the calculus of relations*. Arrow explains to Jerry Kelly what this course was about: "To say it was in the calculus of relations meant that he gave an axiomatic treatment of relations, although he motivated it, of course, by motivating the axioms." (Arrow, Sen, and Suzumura 2011, 5). Suppes (2005) specifies the elements of Tarski's course that were useful in obtaining Arrow's impossibility theorem: "The standard operations on relations covered in the article, such as converse and relative product, were not used by Ken […] It was the clear abstract formulation at an elementary level of ordering relations that had the largest impact." (Suppes 2005, 321). This course introduced Arrow to topics such as transitivity and ordering relations.

[84] This is an extract of one of the interviews conducted by Kristen Monroe and Nicholas Lampros between 2010 and 2013.

In the summer of 1940, Tarski asked him to proofread his textbook, *Introduction to Logic*, in order to correct the English (the book had originally been written in Polish, but it was a German version that was translated into English).[85] Arrow undoubtedly did not limit himself to proofreading the textbook. He understood the axiomatic method proposed by Tarski, which was known in mathematics but not in economics (Hibbard 2011; Suppes 2005). Arrow decided to go to Columbia University for graduate studies, partly because it was one of the few universities in the U.S. where the study of mathematical statistics was strongly developed. Another reason why he went to Columbia was the financial aspect. Originally, Arrow wanted to be a statistician, but there was no Ph.D. in statistics at the time at Columbia. He decided to enroll in mathematics, being the closest to statistics. Arrow needed financial support to continue his studies and asked Professor Harold Hotelling, a statistician in the Economics Department, for a letter of recommendation for a fellowship in mathematics. Hotelling had no influence in the Mathematics Department, but helped Arrow to get a fellowship in economics. Arrow got Hotelling's support because he had attended his mathematical economics course. Arrow had impressed him by solving a problem in economic theory that intrigued him (Arrow, Sen, and Suzumura 2011; Feiwel 1987b; Hibbard 2011).

In the mathematical economics course, Arrow learned "the then not widespread idea that consumers choose commodity vectors as a most preferred point (the ordinalist interpretation), as contrasted with the older view that there was a numerical valued utility function which they maximized." (Arrow 1991 [2014b], 145). Arrow identified this view with the logical notion of an ordering.

Undoubtedly the skills that Arrow acquired thanks to his fascination with mathematical logic were essential to developed "an analytical method which allowed him to treat all conceivable voting schemes simultaneously within one unified conceptual framework." (Suzumura 2002, 10). There were pioneering contributions which specified diverse voting schemes such as the well-known method of simple majority rule (analyzed in detail by Condorcet), the Borda method, the Dodgson method, the Black method, and so forth. The first method that Arrow studied when he wrote his final dissertation was simple majority rule. This is clear from his interview with Jerry Kelly. "One natural method of taking a bunch of $R's$ and putting them

[85] The German philosopher Olaf Helmer was the translator. Arrow found Helmer eight years later at the Rand Corporation. See sub-section 3.2.1.

together would be by pairwise comparisons by majority voting." (Arrow, Sen, and Suzumura 2011, 14).[86]

Arrow realized however that the simple majority rule did not always produce a sharply defined winner. He rediscovered the Condorcet paradox, which he did not know when he started studying voting systems. He decided to study other voting rules with the aim of obtaining a different result. He considered the Borda method but he was not aware that it was Borda. It is important to bear in mind that Arrow was often self-taught, also when studying diverse voting procedures. In the interview with Jerry Kelly, this point about voting rules is clear. Arrow admits that he did not ask anybody in the Rand Corporation about these rules:

> JK. So this experimentation was totally isolated. You didn't ask anybody about rules.
>
> KA. Right, but that also reflected me. It seems to me I was trying at some point to systematically go through all possible rules. I took some examples and then considered related examples […] I did grasp that, at some point in this procedure, part of the point was I was only using information on the alternatives under consideration. But then that struck me as a very natural thing to do (Arrow, Sen, and Suzumura 2011, 15).

But after trying out diverse voting rules, one of the points Arrow stresses is a contradictory result when a set of axioms are imposed on voting systems. This contradictory result is based on an original method proposed by Arrow, which we are going to study starting with his concept of the social welfare function. This function "maps each profile of individual preference orderings into a unique social preference ordering" (Suzumura 2002, 10). It denotes the rule for aggregating each profile of individual preference orderings into a social preference ordering. Formally, Arrow's social welfare function (F) maps each set $\boldsymbol{R} = (R_1, R_2, \dots, R_n)$ of individual orderings over the set Y of all possible social states, where $R_i, (i = 1, 2, \dots, n)$ denotes person $i's$ individual ordering, into a social ordering $R = F(\boldsymbol{R})$.

Suzumura (2002) points out the originality of Arrow's social welfare function or constitution.[87] He supposes a society with only two individuals and six profiles $\{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$ of individual

[86] *R's* refer to preference relations.

[87] Arrow spent the period from December 1951 to August 1952 in Europe on a Social Science Research Council fellowship. He gave an exposition of his then-new work on social choice at the *Institut des Sciences Economiques Appliquées* on June 9, 1952. In the article that was published after the conference, Arrow uses the expression "social choice function" instead of "social welfare function." Arrow justifies this change due to criticism from I.M.D. Little who claimed that "the term "social welfare" would make the reader believe that the ordering of social preferences which results from it constitutes a system of ethical judgments by "society"." (Arrow 1952 [1984b], 50). The term "constitution" appears in the section "Notes on the Theory of Social Choice" in 1963. "I will

orderings over three alternative social states x, y, z. Then each of the following orderings can represent the preference ordering of the two individuals over the three social states:

$$\alpha: x, y, z \quad \beta: x, z, y \quad \gamma: y, x, z \quad \delta: y, z, x \quad \epsilon: z, x, y \quad \zeta: z, y, x$$

Even in the simplest imaginable society proposed by Suzumura (2002), there are 6^{36} Arrovian social welfare functions which must be verified one by one for their democratic legitimacy and for their informational efficiency. This task seems impossible. In order to overcome this difficulty, Arrow imposes a set of axioms that allows him to analyze these 6^{36} social welfare functions all at once. This new methodology leads him to obtain his well-known impossibility theorem, which shows that there is no social welfare function satisfying a set of axioms required for democratic legitimacy and informational efficiency.

It is important to note two essential elements of Arrow's social welfare function: social ordering and individual preferences. With regard to social ordering, there are at least two different interpretations worth studying to understand the novelty of Arrow's approach. First, social ordering can reflect, "the results of comparisons of social states under a procedure or rule adopted by the society to rank social states for the purpose of taking decisions" (Pattanaik 2014, 6). This interpretation suggests that the social welfare function is a decision procedure that the society applies in order to rank distinct social states. Second, social ordering can reflect an individual's social welfare judgments, that is to say, "an individual's ethical judgments about the relative goodness or badness of social states." (Ibid). In this interpretation the individual may belong to the society, may be an outsider of the society or may be a central planner.

Arrow is aware of these two distinct interpretations of social ordering, and he expresses it in his answer to Little's (1952) criticism of his social welfare function.[88] "A welfare judgment requires that some one person is judge; a rule for arriving at social decisions may be agreed upon for reasons of convenience and necessity without its outcomes being treated as evaluations by anyone in particular." (Arrow 1963, 106). Arrow opts for the decision procedure interpretation of social ordering, since he thinks that the results of the decision procedure

therefore now use the term 'constitution,' as suggested by Kemp and Asimakoplos. The difference, however, is largely terminological; to have a social welfare function in Bergson's sense, there must be a constitution" (Arrow 1963, 105).

[88] "We must now make an important distinction which Arrow fails to draw. He calls his function both a social welfare function and a decision-making process." (Little 1952, 427). According to Little, the social welfare function proposed by Arrow must be interpreted as a decision-making process instead of a social welfare function. In addition, it is important to say that Little (1952) and Bergson (1954) think that welfare economics should deal only with social welfare judgments and not with procedures or rules that the society may embrace in order to take decisions.

implemented by the society to aggregate the preference orderings of individuals would determine the final social choice.

As for social orderings, individual preferences play an essential role in Arrow's approach.[89] He affirms that "it is simply assumed that the individual orders all social states by whatever standards he deems relevant." (Arrow 1951 [1963], 17). Arrow focuses on the category of individual preference orderings as the starting point for social decisions. His social welfare function excludes cardinal aspects of individual preferences and interpersonal comparison of preference intensities. "The viewpoint will be taken here that interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the measurability of individual utility" (Ibid, 9). On the other hand, those preferences do not require to be based only on individual tastes concerning their own consumption. "The individual orderings which enter as arguments into the social welfare function as defined here refer to the values of individuals rather than to their tastes" (Ibid, 23).

We now present the five axioms that gives rise to Arrow's famous impossibility theorem:

- U (Unrestricted domain): The domain of the social welfare function comprises all possible $n-tuples$ of individual preferences $\{R_i\}$.

This means that "every individual is free to form and express whatever preference ordering he/she cares to specify, which represents his/her evaluations of the goodness of social states" (Suzumura 2002, 12).

- IIA (Independence of Irrelevant Alternatives): Let $\boldsymbol{R} = (R_1, R_2, \dots, R_n)$ and $\boldsymbol{R'} = (R_1, R_2, \dots, R_n)$ be two sets of individual preferences. If, for all individuals j and all x and y in a subset S of Y, $xR_jy \rightarrow xR'_jy$, then $xRy \leftrightarrow xR'y$ are the same.

What this means is that when Arrow's social welfare function aggregates individual preferences, it has to take each pair of social states separately, without taking into account preferences for social states other than them.

- P (Pareto Principle): For any pair, x, y in Y, $[\forall_j : xP_jy] \rightarrow xPy$. It means that if everybody strictly prefers some social state x to another social state y, then x is socially preferred to y.

[89] It is important to consider Sen's (1977b; 1997a; 2011; 2014) contributions. He analyzes the different senses in which one can study social ordering and individual preferences, in order to understand the informational bases of different social choice procedures.

This axiom is also called weak Pareto principle. The weaker form of the Pareto principle "says less than the usual Pareto criterion since nothing is stated about a case where someone prefers x to y and everyone regards *x* to be at least as good as *y*." (Sen 1970 [2017], 70).

- D (Non-dictatorship): There is no individual j such that for every element in the social welfare function, $\forall x, y \in Y: xP_j y \rightarrow xPy$.
- O (Social Ordering): The social ranking of alternative social states must be transitive and complete.

Note that this last axiom is not formulated by Arrow. But he integrates it in his notion of the social welfare function when he argues that the result of aggregating individual preferences must be an ordering if it reflects a "rational choice-making" (Arrow 1951 [1963], 19).[90]

What Arrow's impossibility theorem (formally, the "General Possibility Theorem") says is that if there are at least three distinct social states and at least two distinct individuals (not an infinite number of individuals), then no social welfare function can satisfy these conditions simultaneously.

In this sub-section, we studied the value judgments made by Arrow in order to choose the method that allows him to obtain his impossibility theorem. We consider that his interest in mathematical logic and voting procedures are essential in the choice of an axiomatic framework, which leads to the birth of social choice theory in its contemporary form. Arrow's definition of individual preferences is related with votes. A person's vote reflects what she considers relevant.

3.2.3 The Choice of the Standards of Validity

We consider that it is evident that the axiomatic approach, which was unfamiliar in economics before Arrow's work, leads Arrow to write out the proof of his impossibility theorem.[91] This proof allows him to obtain "a result of breathtaking elegance and power" which has often been

[90] "Throughout this analysis it will be assumed that individuals are rational […] so that R may also be assumed to satisfy Axiom I and II" (Arrow 1951 [1963], 19). Axiom I: For all x and y, either xRy or yRx (completeness). This axiom holds when $x = y$, as well as $x \neq y$, it means that x is indifferent to itself for any x, and this entails xRx. In this way, a relation satisfying this axiom is both complete and reflexive. Axiom II: For all $x, y, and\ z$, xRy and yRz imply xRz (transitivity).

[91] See the appendix for a proof of Arrow's theorem based on Sen (2011).

interpreted as disastrous with regard to the possibility of having a democratic social choice (Sen 1999, 351).

Arrow was influenced by Bertrand Russell's *Introduction to Mathematical Philosophy* (1920). He read this book while in high school in the 1930s. Arrow says in his interview with Kelly how useful this book was for him: "I learned the ideas of mathematical logic and its applications to mathematics in Russell's book." (Arrow, Sen, and Suzumura 2011, 4). As we have already mentioned in the previous sub-section, Tarski's *Introduction to Logic* (1941) played an important role in Arrow's learning of the axiomatic approach that he introduced in welfare economics. Arrow also learned on his own, reading for example: Birkhoff and MacLane, *A Survey of Modern Algebra* (1941) (Suppes 2005).[92]

In order to prove his impossibility theorem, Arrow (1951 [1963]) starts by analyzing the method of majority decision. He concludes that if there are only two alternative social states, this voting procedure satisfies all the conditions of his theorem. In this way, he demonstrates that when the formation of a social welfare function is given by two individuals who express their preferences for three alternative social states, this social welfare function does not satisfy all the conditions simultaneously.

Once Arrow demonstrates that the method of majority decision does not satisfy all the conditions of his theorem, he has to see whether this result can be generalized to other voting systems. He shows that, "if no prior assumptions are made about the nature of individual orderings, there is no method of voting which will remove the paradox of voting […], neither plurality voting nor any scheme of proportional representation." (Arrow 1951 [1963], 59). It is important to note that the interpretation given by Arrow of the result of the impossibility theorem highlights that "the social ordering be formed from individual orderings and that the social decision between two alternatives be independent of the desires of individuals involving any alternatives other than the given two […]" (Ibid). This implies that conditions U and IIA taken together lead to excluding the interpersonal comparisons of utility from the analysis proposed by Arrow. Thus, Arrow's social welfare function, that we have already defined as a decision procedure, does not allow the use of interpersonal comparisons of utility. It means that the utility information that is used in this structure of social choice is made up of n-tuples of individual preferences that are considered separately (Sen 2011).

[92] Arrow (1951 [1963]) refers to Tarski's *Introduction of logic* in two footnotes on p. 13 and in one footnote on p. 14, in order to make clear his terminology about ordering relations.

Arrow is aware of the importance of his contribution and of the implications of his impossibility theorem. "When I got the result, I felt it was significant […]. This was at least asking some very fundamental questions about the whole nature of social intercourse and particularly about legitimation of collective action." (Arrow, Sen, and Suzumura 2011, 16). He proposes to relax some of the conditions of his theorem in order to make social welfare judgments that are not dictatorial. He provided some insight to overcome this result and argued that the condition that has to be relaxed is U. Arrow dedicates two chapters (Chapter VI and Chapter VII) of his SCIV (1963) to analyze the different possibilities to restrict the wide range of sets of individual orderings that form a social ordering. In other words, it supposes restricting the possible preference profiles of individuals with the aim of defining better the similarity of evaluations within the same society (Arrow 1951 [1963], 60).

So in this section we analyzed the methodological value judgments of Arrow which led him to develop the social choice theory. "When I decided I really had a dissertation topic, the Social Choice, there wasn't anybody in the field. The field didn't exist." (Hibbard 2011, 90). We can say that the academic context surrounding Arrow (at the City College of New York, Columbia University, the Cowles Commission and the Rand Corporation), and his own interest in mathematical logic and in election systems were all essential to his inquiry in social choice theory.

3.3 The Value-ladenness of Arrow's Impossibility Theorem

In this section, we study the value-laden aspect of Arrow's impossibility theorem. We want to show that value judgments do not only influence the choice of the topic and the goals of Arrow, but they have an impact on the interpretation of his seminal work. We consider the neutrality theses proposed by Mongin (2006b) to determine what position Arrow assumed regarding the role of value judgments in his impossibility theorem. We have already analyzed the four theses to study the value neutrality problem in the first chapter of this dissertation. We now focus on arguing which of these theses fit better into Arrow's impossibility theorem. Mongin (2006b) provides an answer himself to our questioning. We study Mongin's arguments to establish the thesis that fits better into Arrow's impossibility theorem.

3.3.1 Analyzing Mongin's Hypothesis: Arrow is a Weak Neutrality Theorist

Mongin considers that Arrow expresses a weak neutrality position (thesis 2 in Mongin's paper) regarding his impossibility theorem. There is the assumption that "the question of making value judgments logically arises for the economist *qua economist* […] There are occasions in which he might, and even should, make these judgments" (Mongin 2006b, 261).

Mongin identified the new welfare economics with his thesis 2. This is clear in the following quote:

> The new welfare economists felt that they would not employ the Paretian apparatus, even *qua economists*, unless they would invest it with 'ethical' interest. Thus, they *added* to the definition a separable value judgment to connect the Pareto ordering with the 'economic' or 'social' good (Mongin 2006b, 260).

In other words, the new welfare economists assume that they can separate logically and practically value judgments from judgments of fact. This is one of the three premises of what Mongin calls the containment claim. The other two premises establish that the value judgments made by economists are relatively few in number and that they are easy to recognize. According to Mongin, the containment claim is essential in the definition of the weak neutrality thesis.

Mongin not only considers that the new welfare economics illustrates his thesis 2, but he also argues that Arrow in the development of the impossibility theorem can be considered as a weak neutrality theorist. Mongin (2006b) mentions Arrow's position regarding value judgments explicitly in three passages of his article. We analyze them in order to understand why Mongin identifies Arrow's impossibility theorem with the weak neutrality position.

The first passage is the following: "He [Arrow] applies to this novel condition [IIA] the two-part analysis they [new welfare economists] apply to the Pareto principle; i.e., he [Arrow] keeps his definition of IIA carefully separate from his value judgment that rational groups should satisfy it" (Mongin 2006b, 260). In other words, Mongin refers to the implications of condition IIA on Arrow's social welfare function: "The collective choice made from a given set of options should be invariant with respect to changes in the individual preferences concerning options outside the set" (Arrow 1952 [1984b], 51). Then Arrow adds the value judgment that society must satisfy this condition to connect individual preferences for relevant alternatives with the collective choice.

The second passage is a footnote in which Mongin explicitly says that Arrow takes sides on the role of value judgments in his work. "Mongin (1999, 2000) investigates Arrow's work with a view to showing that he does not refrain from taking sides, contrary to the popular view of social choice theory as mostly formalistic" (Mongin 2006b, 284, footnote n° 4). This quote reveals that Mongin is aware that considering Arrow as a weak neutrality theorist is not a view shared by social choice theorists. A social choice theorist can be tempted to identify Arrow's position regarding value judgments with Mongin's strong neutrality thesis. This thesis assumes that "in their professional work economists should always refrain from making value judgments" (Mongin 2006b, 259). It means that economics neither is nor ought to be normative. It is clear that placing Arrow in one or another value neutrality thesis cannot be taken for granted. For this reason, we will analyze what Mongin (1999, 2002) says about value judgments and Arrow's impossibility theorem. This will clarify the reasoning that led Mongin to identify Arrow's work with his thesis 2.[93]

The third passage that shows Arrow's position regarding value judgments in Mongin's work is the following: "Weak neutrality theorists, such as Bergson, Samuelson and Arrow, all recognized that value judgments were needed to give the social welfare function a content, but only Arrow let the economist *qua economist* carry the task of filling out this content" (Mongin 2006b, 282). Mongin refers to the importance that Arrow gives to the values of all individuals in order to make social welfare judgments. The following quote clearly shows how Arrow *qua economist* considers value judgments as necessary to fill out the content of the social welfare function:

> We can suppose that the individuals hold value judgments on the method of formation of collective decisions; the aggregation of the ethical feelings of individuals is itself an object of ethical judgments. The conditions of which we speak form a set of these value judgments that can be regarded as unanimously acceptable (Arrow 1952 [1984b], 51).

Behind Arrow's position, highlighted by Mongin, regarding the content of the social welfare function, there is the controversy about the Bergson-Samuelson social welfare function as a consequence of Arrow's impossibility theorem.[94] This controversy confronts welfare

[93] Mongin (2000) was presented at the ESTHET 2000 Conference (Universität Graz). However, it was published in 2002 in *Is There Progress in Economics? Knowledge, Truth and History of Economic Thought*, edited by Stephan Boehm, Christian Gehrke, Heinz Kurz, and Richard Sturn. For this reason, we refer to the 2002 publication and not to his lecture given in 2000 both in the quotations and in the references.

[94] We refer to the individualistic social welfare function as defined by Samuelson (1981). This function "ethically orders the various states of the world" and "lets individual tastes 'count,' in the sense of agreeing with individuals'

economists and social choice theorists. It has been studied in an analytical and technical way by Mongin (2002) and Fleurbaey and Mongin (2005), and from the viewpoint of economic history by Igersheim (2017).

We focus on one aspect of this controversy that can be useful to understand why Mongin (2006b) considers that only Arrow let the economist *qua economist* to fill out the content of the social welfare function with value judgments. We refer to a major point made by Little (1952) and Bergson (1954) that a Bergson-Samuelson social welfare function does not reflect the value judgments of all members of society, but only the value judgments of an individual, who ordered the states of society. "The person will communicate his evaluative judgments to the welfare economist, who should be able to summarize them into a coherent criterion, that is, an ordering" (Mongin 2002, 154). But the construction of this ordering is not a trivial step, and Arrow's conditions show the importance of knowing how to construct a welfare ordering for an individual.

In order to illustrate this point, we present an example given by Arrow (1952 [1984b]) in which he takes an individual named Primus who proposes a new ordering based on both his ethical viewpoint and the preferences of the different individuals of the society. Arrow considers that the conditions of his theorem are "value judgments which might reasonably be part of the ethics of Primus" (Ibid, 56). He assumes *qua economist* that value judgments are necessary to fill out the content of his social welfare function. The aim however of this example is to show that Primus "will find […] that he cannot construct a preference scale which would express his ethical choices among various distributions and would be in accord with these value judgments" (Ibid). Although Arrow's aim with the Primus example is to show that his impossibility result affects the Bergson-Samuelson social welfare function, we think that this example also shows that Arrow considers the conditions of his theorem as value judgments.

Even if the three passages we analyzed seem to indicate that Arrow can be identified with the weak neutrality thesis regarding his impossibility theorem, we consider that limiting ourselves to these passages is not enough to identify the role of value judgments in Arrow's impossibility theorem. For this reason, we now analyze two previous articles by Mongin (1999, 2002) where he establishes the bases that lead him in 2006b to place Arrow's work in his thesis 2. It is the

orderings when those orderings are unanimous and resolving them ethically when they are not unanimous" (Samuelson 1981, 224)

interconnection between Mongin's three articles that will allow us to understand why he considers Arrow as a weak neutrality theorist.

3.3.2 Three Ways of Interpreting the Role of Value Judgments in Arrow's Impossibility Theorem

Mongin (1999) proposes three ways of interpreting the role of value judgments in Arrow's impossibility theorem. First, Mongin argues that Arrow does not make value judgments by himself, but rather he conducts a logical analysis of the consistency of different value judgments in mechanisms of social choice. Second, Mongin considers that Arrow makes value judgments by himself when it is appropriate to analyze the conditions of his theorem. Third, Mongin claims that when Arrow is aware of the result of his theorem, he proposes to relax some of the conditions imposed. It implies that value judgments play an important role in the choice of the condition which is relaxed.

The importance given by Arrow to the coherence of diverse value judgments in any mechanisms of social choice can be seen in the following quote:

> The economist should investigate those mechanisms for social choice which satisfy the value judgments and should check their consequences to see if still other value judgments might be violated. In particular, he should ask the question whether or not the value judgments are consistent with each other, i.e., do there exist any mechanisms of social choice which will in fact satisfy the value judgments made? (Arrow 1951 [1963], 4–5).

According to Mongin (1999), Arrow's position on value judgments can be examined by studying his criticism of new welfare economics. One the one hand, when Arrow examines the compensation principle proposed by N. Kaldor and J.R. Hicks in 1939, he shows that the Pareto criterion can only be considered as "a useful preliminary to the analysis of social welfare function," because this criterion does not allow one to determine a complete enumeration of the social ordering (Arrow 1951 [1963], 37). The Pareto criterion conflicts with other value judgments implicit in the social choice function: U, IIA and D (Mongin 1999, 543). Mongin (2002) shows how Arrow implicitly refutes the compensation principle based on his impossibility theorem:

> Take any binary relation R having the following two properties: first, it extends the partial ordering implied by the Pareto principle; second, it is complete. If R is

> obtained from a 'social welfare function,' then assuming the Arrovian conditions other than social ordering, we conclude that R must be intransitive (Mongin 2002, 157).

This argument shows that the value judgments defining the compensation principle are inconsistent with a rational social choice. It means that Pareto superiority, which is the point of departure of the compensation principle, is considered as insufficient to determine a complete ordering of social states.

On the other hand, we must consider Arrow's position regarding the Bergson-Samuelson social welfare function. We will not make an exhaustive study of the controversy that arises regarding the conceptual differences between the Arrow social welfare function and the Bergson-Samuelson social welfare function. Instead, we focus on analyzing how Arrow illustrates that his impossibility result affects the Bergson-Samuelson social welfare function. This allows us to examine if Arrow conducts a logical analysis when he studies the consistency of the Bergson-Samuelson social welfare function.

The following quote allows us to see Arrow's early analysis of the Bergson-Samuelson social welfare function:

> The Bergson social welfare function is mathematically isomorphic to the social welfare function under individualistic assumptions. Hence the Possibility Theorem under Individualistic Assumptions […] is applicable here; we cannot construct a Bergson social welfare function […] that will satisfy Conditions 2-5 and that will lead to a true social ordering for every set of individual tastes (Arrow 1951 [1963], 72).

This quote is at the end of chapter 6 of Arrow's book. In this chapter he proposes a solution to his impossibility theorem based on the modification of condition U. He supposes that "some but not all choices made by an individual are known in advance" (Ibid, 61). It means that Arrow restricts the wide range of sets of individual ordering of his original condition U. He imposes two conditions on the individual preferences orderings, that is, the individualistic assumptions. The first condition establishes that "each individual's comparison of two alternative social states depends only on the commodities that he receives […] in the two states." The second condition establishes that "in comparing two personal situations in one of which he receives at least as much of each commodity […] and more of at least one commodity than in the other, the individual will prefer the first situation" (Ibid). He demonstrates however that in a world of more than one commodity, these restrictions do not prevent that the impossibility result reappears.

Arrow thus shows that his attempt to overcome the impossibility theorem is not possible by restricting the condition U with the imposition of individualistic assumptions (We will refer to this restricted domain assumption as condition U'). This result allows Arrow to claim that the Bergson-Samuelson social welfare function is affected by his "Possibility Theorem for Individualistic Assumptions" (Ibid, 63). Note that Mongin (2002) considers that the way Arrow concludes that the construction of this function is affected by this Possibility Theorem is a "purely logical step" (Mongin 2002, 153).The problematic step is to vindicate each of conditions (U', IIA, P, D and O) "in terms of the general objective and privileged interpretations of welfare economics" (Ibid). Mongin considers that the way in which Arrow defends each of his conditions is not enough to convince welfare economists that his impossibility result affects the Bergson-Samuelson social welfare function. "Arrow (1963: 73) is disappointingly brief when it comes to this fundamental discussion. Basically, he contents himself with pointing out again the general plausibility of his conditions" (Ibid).

For this reason, Mongin (1999) proposes that one way in which Arrow can defend his conditions is by making value judgments himself, which allow him to justify his position regarding the Bergson-Samuelson social welfare function. Arrow argues that the five conditions of his theorem are value judgments and could therefore be questioned. "Taken together they express the doctrines of citizens' sovereignty and rationality in a very general form, with the citizens being allowed to have a wide range of values." (Arrow 1951 [1963], 31). Arrow is aware of the essential role of individual value judgments in the formation of collective decisions. "The individual plays a central role in social choice as the judge of alternative social actions according to his own standards. We presume that each individual has some way of ranking social actions according to his preferences for their consequences" (Arrow 1967 [1984c], 66). These preferences form an individual's value system, and they are represented by an ordering.

Mongin (1999) emphasizes two ways in which Arrow makes his own value judgments. First, Arrow considers that the Pareto principle (P), the non-dictatorship condition (D) and the unrestricted domain condition (U) express an idea of individuals' sovereignty in collective well-being. But Arrow considers explicitly that the independence of irrelevant alternatives (IIA) and the social ordering condition (O) constitute conditions of rationality.

> It seems reasonable to say that the minimum of rationality that one can demand of someone who must make decisions is that his choices be in conformity with an ordering or a scale of preferences endowed with the properties explained above […] An individual is rational if his preferences among candidates can be expressed by an ordering; similarly, collective decisions are made rationally if they are

> determined by an ordering acceptable to the entire society (Arrow 1952 [1984b], 49).

"The properties explained above" are transitivity, completeness, and reflexivity. This quote makes it clear that condition O constitutes a condition of rationality. The condition IIA constitutes a condition of rationality because it implies that the choice made by society from a given environment does not consider individual preferences among options that are not in fact available in the given environment (Arrow 1967 [1984c], 62–3) .

Considering that conditions P and D express individuals' sovereignty in collective well-being might seem less evident than in the case of condition U. The latter condition claims explicitly that each member of society can freely express their preferences, whatever they are. It means that the social welfare function should be capable of taking into account a wide range of individual preferences in order to construct an ordering for the collectivity.

Regarding condition P, it is important to note that in the proof of the impossibility theorem proposed in the 1963 edition, Arrow replaces some conditions by condition P. These conditions are the positive association condition and the condition of a citizen's sovereignty. We consider that analyzing these two conditions allow us to better understand why condition P reflects individuals' sovereignty in collective well-being.

On the one hand, the condition of positive association states that "if one alternative social state rises or remains still in the ordering of every individual without any other change in those orderings, we expect that it rises, or at least does not fall, in the social ordering" (Arrow 1951 [1963], 25). Arrow claims that this condition "bears on the relation between the individual orderings and the collective ordering which result from them: the collective ordering should reflect positively the individual welfare judgments" (Arrow 1952 [1984b], 51).

On the other hand, the condition of a citizen's sovereignty captures the idea that the "social welfare function is not to be imposed" (Arrow 1951 [1963], 29), which means that the social welfare function depends on individuals' preferences. Amadae (2003) emphasizes the importance of a citizen's sovereignty in Arrow's theorem: "Arrow requires that legitimate collective decisions be based on individuals' preferences and, by virtue of his commitment to the notion of citizens' sovereignty, upholds the individual as the final arbiter of his own preferences" (Amadae 2003, 110–11). While Arrow claims that both conditions correspond to the definition of a consumer's sovereignty, but since he refers to values rather than tastes, he

claims that it is more appropriate to refer to such conditions as representing the idea of citizens' sovereignty.

Regarding condition D, Arrow claims that a dictatorial social welfare function should be excluded if "we are interested in the construction of the collective methods of social choice" (Arrow 1951 [1963], 30). This condition reflects the individuals' sovereignty in collective well-being, since it does not allow social choices to be based on the preferences of a single individual.

Mongin (1999) claims that the other way Arrow makes value judgments is with respect to the position he takes on interpersonal comparisons of utility. Arrow explicitly expresses:

> It seems to make no sense to add the utility of one individual, a psychic magnitude in his mind, with the utility of another individual […] The behavior of an individual in making choices is describable by means of a preference scale without any cardinal significance, either individual or interpersonal (Arrow 1951 [1963], 11).

Arrow assumes ordinal preferences without interpersonal comparisons of utility in his early analysis. It is not until the publication of two articles in 1977, that Arrow is concerned with the explaining in more detail of the controversy between the cardinal and ordinal significance of preferences and of the possibility and meaning of interpersonal comparisons of utility. He (1977a) emphasizes that the separation between cardinality and interpersonal comparability is logical. It implies that each individual might have cardinal preferences, but the latter are not necessarily interpersonally comparable. He considers that interpersonal comparisons of preference intensities (or cardinal-difference comparisons) are not meaningful "partly on the grounds that the implied cardinality of individual preference has no verifiable meaning" (Arrow 1977a, 613).

Arrow then proposes to make interpersonal ordinal comparisons based on the principle of extended sympathy, which allows him to consider as significant statements of the form: "individual i in state x is better off than individual j in state y" (Arrow 1977b, 220). Interestingly, he has already mentioned the principle of extended sympathy in his book (Arrow 1963, 114–15), but it is in his 1977b article that he makes a formal analysis of the meaning of interpersonal ordinal comparisons.[95] He concludes however that this form of interpersonal

[95] Arrow (1977b) proposes a way to formally interpret interpersonal ordinal comparisons:
"We may suppose a space, Y, which defines the range of possible implications of a social state for an individual. Since the state defines, for each individual, everything that characterizes his satisfactions, the space Y is the same for all individuals. It includes goods, tastes, and the reactions of others to the extent that individuals care about each other. All individuals have the same preferences over Y. Let $u(y)$ be the ordinally defined utility indicator. Each state in X defines implications for every individual. Let $G_i(x)(i = 1, \dots, n)$ be for each individual a mapping

comparability denies the autonomy of individuals since it reduces the individual to a specified list of qualities, which affects his individuality. "No doubt it is some such feeling as this that has made me so reluctant to shift from pure ordinalism, despite my desire to seek a basis for a theory of justice" (Arrow 1977b, 225).[96]

Regarding the third interpretation of the role of value judgments in Arrow's impossibility theorem proposed by Mongin (1999), it is important to note that he focuses on highlighting Arrow's theoretical inclination to relax condition U. This is one of Arrow's (1963) proposals to overcome his impossibility result. He restricts the possible types of individual preferences: "It is required to find a social welfare function which will be satisfactory for all sets of individual values compatible with those restrictions but not necessarily satisfactory or even defined for other types of values" (Arrow 1951 [1963], 60). In this way, individuals' choices are determined by limited sets of values. This leads Arrow to assume that the shape of an individual's preferences cannot be known in advance, but he assumes that individuals share common values, which are the basis of a homogeneous society.

Arrow (1951[1963]) analyzes the restriction of single-peaked preferences proposed by Professor Duncan Black in 1948. An example in where this restriction is satisfied is the left-right ordering of political parties. This can be considered as "a linear axis, independent of the voters, on which alternatives are positioned." (Escoffier, Lang, and Öztürk 2008, 2). A voter has a single-peaked preference with respect the left-right axis if, on each side of the "peak" (i.e., the preferred candidate), her preference grows with the closeness to the preferred candidate. In this way, assuming the condition of single-peaked preferences, the method of majority decision will deliver an ordering, if there is an odd number of voters. This last requirement is considered peculiar or even severe, since it limits the operational significance of the possibility theorem (Sen 1969).[97]

Arrow's aim in examining the case of single-peaked preferences is to make clear the importance of a similarity of attitudes in the formation of social judgments.

---

from X to Y, expressing these implications. Then we can identify, $u(x,i) = u[G_i(x)]$, and this possesses Co-ordinal Invariance" (Ibid, 224).
"Co-ordinal Invariance: If there exists a strictly monotone function, g from real numbers to real numbers such that, $u'(x,i) = g[(x,i)]$, all x and i, then $f(u) = f(u')$" (Ibid, 220).

[96] This reference to the theory of justice is based on the relationship between social choice theory and Rawls's maximin criterion studied by Arrow (1973; 1977a; 1977b).

[97] There are of course other conditions to restrict the field of admissible preferences, for example: "Inada's condition of 'single-caved preferences' and 'preferences separable into two groups,' Ward's condition of 'Latin-square-less preferences,' and Sen's condition of 'value restricted preferences'." (Sen 1969, 389).

> The fact that Black's restrictions on individual orderings suffice to permit a social welfare function casts new light on what is meant by similarity of social attitudes […] individuals can have varied first choices; but they must have a fundamentally similar attitude toward the classification of the alternatives since they all order the alternatives in the same way (Arrow 1951 [1963], 80).

The assumption of single-peakedness implies that individual preferences may have special structures that lead to restrict the domain of admissible set of individual ordering relations. Arrow proposes another way to analyze the assumption of complete unanimity in order to relax condition U. He examines the proposition of the idealist school, which is represented by philosophers such as Kant, Rousseau, T.H Green, among others. Arrow considers that the concept of "general will" implies a unanimous agreement that is the basis of social morality. He summarizes the position of the idealistic doctrine by saying that individuals have two orderings, one that reflects their everyday actions (tastes), and the other their "moral imperative." If the latter ordering is considered relevant to social choice, Arrow assumes that a complete unanimity may rise.

Significantly, Arrow tries to elucidate the basis of a consensus of the moral imperative when discussing the debate between Mr. Dobb and Professor Lerner. Dobb questions the reliability of the individual will, which can be corrupted by "the variability of individual tastes in the presence of advertising and of the emergence of new products." Arrow says that Lerner is right when expressing that Dobb's position "implies some transcendental optimum other than that shown 'by a free market or in any other way'" (Arrow 1951 [1963], 84). This means that the consensus of the moral imperative would be achieved by transcending individuals' tastes. Arrow considers that this would imply an ethical absolutism, which transgresses liberalism. He argues that the consensus of the moral imperative can be based on individual preferences. [98]

This unanimity as the basis of social welfare judgments cannot be achieved through the market mechanism which cannot consider, "the altruistic motives which must be present to secure that consensus." He adds that "this does not deny the possibility of a limited use of the market as an instrument for achieving certain parts of the social optimum" (Arrow 1951 [1963], 86).

[98] Amadae (2003) and Saito (2011) discuss Arrow's position regarding the idealist school. Amadae argues that Arrow's theorem "rejects the possibility that a social consensus on ends could emerge as a result of a philosophical ideal transcending individual's desires as a guide to collective decisionmaking" (Amadae 2003, 114). Saito analyzes the debate between Dobb and Lerner proposed by Arrow in order to show that "unlike idealistic scholars who distinguish revealed from ideal preferences, Arrow argues that moral rules can be based on revealed preferences" (Saito 2011, 14).

The question that we can ask ourselves after the analysis proposed by Arrow is what mechanism allows one taking into account the altruistic motives that give rise to the similarity of social attitudes. Berthonnet and Desreumaux (2014), think that the homogeneous values of society are determined outside the market and that voting is the coordination mechanism. Voting takes account of altruistic motives which bring about a similarity of attitudes toward social alternatives. It implies that voting must make it possible to establish common values that allow individual orderings of social alternatives to be reflected in the market. "The market will be able to rationally aggregate individual choices that do not relate directly to values, since the choices of values have already been made in the society by voting" (Berthonnet and Desreumaux 2014, 118).

The analysis proposed by Berthonnet and Desreumaux supposes that Arrow modifies his initial assumption that there is no difference between voting and the market mechanism in the study of social choice. Voting and the market mechanism become two coordination mechanisms that complement each other.[99] This interpretation somehow shows us the importance of the difference that Arrow establishes between tastes and values. Arrow explicitly considers that the market mechanism takes into account only the ordering according to the consumption of the individual (tastes), which implies that the ordering according to the general standard of equity of the individual (values) are not considered by the market mechanism.

To sum up, we have shown throughout the analysis of Arrow's impossibility theory under the prism of Mongin that the weak neutrality thesis is the most appropriate to interpret Arrow's seminal contribution. We consider that we can neither rule out any of the interpretations proposed by Mongin (1999) nor favor one more than another. These interpretations are complementary and therefore have to be analyzed together in order to understand why Mongin (2006b) considers Arrow as a weak neutrality theorist. So the first interpretation assumes that it is logically possible for Arrow *qua economist* to make value judgments and the other two interpretations show how Arrow might, and even should, make value judgments.

[99] See also Amadae (2003), who sees the importance of Arrow's work for the capitalist democracy in the context of the Cold War. She stresses that Arrow's definition of "citizens' sovereignty" implies that the individual is the final arbiter of his personal preferences, which is only respected in Western liberal democracies. "Arrow's condition of citizens' sovereignty is presented as the philosophical alternative to elite rule by an authoritarian government that claims to know what is best for the people" (Amadae 2003, 116).

3.4 Escaping From Arrow's Impossibility Theorem

Beyond Arrow's proposal to relax condition U, there are other alternatives in order to escape from the impossibility result. These include, among others, relaxing the demand that all social preferences be transitive.[100] Loosening condition IIA allows "social welfare functions to make use of more of the information that is carried by individual preference orderings" (Morreau 2019). Another way to overcome the impossibility result is to extend Arrow's framework with the aim to allow individuals to give richer information than is carried by preference orderings. On the one hand, we can consider adding more ordinal information about individual preferences than is allowable within Arrow's approach. On the other hand, cardinal information can be considered together with interpersonal comparisons of utilities (Morreau 2019).[101]

In this section we evaluate Arrow's propositions to escape from his impossibility theorem in light of the assumption that he is a weak neutrality theorist. On the one hand, Arrow had already considered relaxing condition U in his 1950 paper ("A Difficulty in the Concept of Social Welfare"). On the other hand, in chapter VIII of the second edition of SCIV (1963) he considered his condition IIA to be very strong and suggested that interpersonal ordinal comparisons be allowed.

One of the premises of the containment claim (sub-section 3.3.1) is that weak neutrality theorists can separate value judgments from the judgments of fact. We think that Arrow makes a judgment of fact when he suggests that the relaxation of condition U is grounded on the formation of social welfare functions where there tend to be a similarity of attitudes toward social alternatives (Arrow 1951 [1963], 69). He identifies the importance of the value system of individuals for social actions in order to relax condition U. "I want to emphasize that value judgments in favor of limited social preference, just as other value judgments emphasizing social solidarity, must be counted as part of the value systems which individuals use in the judgments of alternative social actions"(Arrow 1967 [1984c], 67). Then Arrow separates the

[100] Sen (1969) shows that a slight relaxation of the transitivity requirement, by demanding transitivity only of strict preference (quasi-transitivity: $xPy \,\&\, yPz \;\rightarrow xPz$), invalidates Arrow's result. However, Allan Gibbard "demonstrated immediately that new problems arise with the weakening of the transitivity condition, because quasi-transitivity yields an oligarchy (making a group of people jointly decisive with veto power for each member of that privileged group)." (Sen 2017, 273). See also Salles (1976), Bordes and Salles (1978), among others.

[101] However, "admitting cardinality of utilities without interpersonal comparisons does not change Arrow's impossibility theorem at all, which can be readily extended to cardinal measurability of utilities […] In contrast even ordinal interpersonal comparisons is adequate to break the exact impossibility." (Sen 1999, 357).

fact that there are values that lead to the similarity of social attitudes from the value judgment that society should adopt some values more than others.

Considering Arrow as a weak neutrality theorist can explain why he proposes to accept interpersonal ordinal comparisons. After interpreting his impossibility theorem in *A Difficulty in the Concept of Social Welfare*, Arrow makes it clear that he is not willing to allow interpersonal comparison of utilities to solve the impossibility result. "It will continue to be maintained that there is no meaningful interpersonal comparison of utilities […]" (Arrow 1950 [1984a], 24). This position leads us to assume that Arrow does not intend to relax condition *IIA*, which guarantees the ordinality of individual preferences and their non-comparability. However, Arrow recognizes that this condition is strict and proposes to accept one type of use of irrelevant alternatives.[102] He thus refers to the principle of extended sympathy already mentioned in the previous section. The following quote reveals Arrow's position:

> I now feel, however, that the austerity imposed by this condition is stricter than desirable. In many situations we do have information on preferences for nonfeasible alternatives. It can certainly be argued that when available this information should be used in social choice […] The potential usefulness of irrelevant alternatives is that they may permit empirically meaningful interpersonal comparisons (Arrow 1967 [1984c], 76).

In this way, Arrow decides to make interpersonal comparisons of the extended preferences type, which allows comparisons of the kind: "State x is better (or worse) for me than state y for you" (Arrow 1963, 114). He notes that the principle of extended sympathy is consistent with the assumption that all value judgments are at least representable as hypothetical choices among alternatives options. Arrow admits the meaningfulness of interpersonal ordinal comparisons as a suggestion to avoid his impossibility theorem. It is important to note that Arrow *qua economist* continues to defend his ordinalist position but accepts interpersonal comparisons. It implies that he continues to regard interpersonal cardinal comparisons as unverifiable since it is about the intensity of preferences. "Cardinal-difference comparisons are not allowed as meaningful, partly on the grounds that the implied cardinality preference of individual preference has no verifiable meaning." (Arrow 1977a, 613).

It is important to highlight Arrow's position more than thirty years later, after the first publication of SCIV (1951), regarding the loosening of one of the conditions of his

[102] Pazner (1979) points out that the principle of extended sympathy can be considered as "irrelevant" because it is not feasible to be oneself and someone else.

impossibility theorem. He clearly shows his interest in the relaxation of condition IIA in his interview with Professor Kelly (1986):

> Well, if I had to pick just one, it would be reformulating a weakened form of the independence of irrelevant alternatives which stops short of just dropping it completely […] I don't want to be in position of saying, well the whole thing depends on the whole preference ordering. My current feeling is that that is the most central issue—the most likely way of really understanding issues (Arrow, Sen, and Suzumura 2011, 23).

Condition IIA has two aspects. First, the *irrelevant* aspect, requiring that only properties of the alternatives in a subset S are considered in determining social preferences over S. Second, the *ordering* aspect, which implies that individual orderings must be both ordinal and not interpersonally comparable at the same time. Relaxing this condition implies violating one or both aspects (Sen 1970 [2017]; Hammond 1987).[103]

One way to relax condition IIA consists of adding more ordinal information about individual preferences in order to overcome the impossibility theorem. This can be presented as a decision mechanism that score or grade the alternatives that individuals prefer, which allows one to aggregate graded inputs that are not available for preference orderings.[104] In an interview with Aaron Hamlin, the co-founder of The Center of Election Science in the U.S., Arrow affirms his preference for a score mechanism:

> Well, I'm a little inclined to think that score systems where you categorize in maybe three or four classes […] is probably the best […] In France, [Michel] Balinski has done some studies of this kind which seem to give some support to this scoring method (Hamlin 2012).

Later on Arrow (2014a) confirms his point of view but shows his doubts regarding the decision mechanism to choose if the condition IIA is dropped:

> But by dropping the independence of irrelevant alternatives condition, you could then employ Borda count, the method proposed by Balinski and Laraki, or perhaps some other alternative. Regardless, what emerges is a consistent ranking, which satisfies the other conditions, not based on cardinal utilities but instead on the rankings […] Whether this is a good decision mechanism or not, I am not prepared to say (Arrow 2014a, 59–60).

[103] The irrelevant aspect of IIA is violated "when, in the social choice involving x and y, the individual rankings of a third alternative, say z, vis-à-vis either x or y or any other alternative, become a relevant factor, with an influence." The ordering aspect is violated "if in the social choice involving x and y anything other than the individual orderings over x and y get a place, e.g., preference intensities." (Sen 1970 [2017], 142–43).

[104] See Balinski and Laraki (2007) who showed that scoring or grading allows one to escape from Arrow's impossibility theorem while staying within an ordinal approach.

Arrow seems to accept the loosening of the *irrelevant* aspect of IIA. This could imply using a weaker condition, such as the one proposed by Pazner (1979).[105] The Independence of Non-Indifference Alternatives ($INIA$) allows the social comparison of two alternatives (x, y) to depend on their utility levels and on how they are ranked by individuals considering other alternatives. Formally, using the notation $C_i(x, x)$ for the indifference set of x, and considering the following clarifications: $\vec{R}$ denotes a profile of individual orderings by $\vec{R} = (R_1, \dots R_n)$ or $\overrightarrow{R'} = (R'_1, \dots R'_n)$. Let a social ordering function F be a mapping from some domain $D^F \subseteq (P_X)^n$ of profiles of individual ordering to P_X. Let P_X be the set of all orderings on X.

***INIA*:** $\forall\, \vec{R}, \vec{R}' \in D^F, \forall x, y \in X, [\big(C_i(x,x) = C_i'(x,x)\big)\ \&\ (C_i(y,y) = C_i'(y,y), \forall_i \in N] \Rightarrow (xRy \Leftrightarrow xR'y)$[106]

Nevertheless, Arrow does not take a position at any time regarding what kind of loosening this axiom must undergo to get out of the impossibility result. In fact, we can say that he not only seems to accept to relax the *irrelevant* aspect of IIA but also the *ordering* aspect of IIA. In the following quote, Arrow affirms that it is important to continue exploring to develop an interpersonal scale:

> If you look at the questionnaires that economists use to assess, for instance, the happiness of people, they will ask a respondent, "How happy are you?" The person then ranks their happiness on scale. Now a rock bottom, hard-boiled economist might say to the individual, "You had happiness of one, and now it's three? What do you mean?" (Arrow 2014a, 60).

Arrow assumes that if people answer these kinds of questions, it is because they consider them meaningful. In addition, the answers of individuals allow a systematic representation. Arrow recognizes that he does not have a mathematical representation of an interpersonal scale, but he encourages economists to find a formulation that replaces IIA.

Sen (1970 [2017]) proposes the notion of social welfare functional ($SWFL$), which allows one to map from profiles of individual utility functions to social preference orderings. It means that Sen extends Arrow's framework by considering the preferences of individuals through utility functions. Sen studies diverse ways to consider the measurability and interpersonal

[105] Fleurbaey proposes a weaker axiom than Arrow's axiom (IIA). The Independence of Other Alternatives (IOA) allows that the social comparisons of two alternatives (x, y) to depend only on their utility levels. Then, IOA is more restrictive than $INIA$. He argues that "the Borda rule, which applies the utilitarian criterion to individual Borda scores […] does satisfy […] $INIA$, but not IOA" (Fleurbaey 2003, 357).

[106] This formalization of $INIA$ is based on Fleurbaey and Mongin (2005).

comparability of utilities. He is aware of the problem of arbitrariness of any cardinal scale in choices over a set of alternatives. However, he considers that it is not decisive to say that the cardinal scale is not relevant for social choice. In this way, Sen shows that more than one method of cardinalization can be relevant to collective choice.

Sen (1970 [2017], Chapter 8) makes it clear that it is the exclusion of interpersonal comparisons that leads to a dictatorship. The impossibility result persists even if individuals have cardinal utility functions combined with non-comparability. In his analysis of interpersonal comparisons and the independence of irrelevant utilities, Hammond (1987) states that the difference between the social welfare functional proposed by Sen and Arrow's social welfare function comes to light because individuals' utility functions may be cardinal and/or interpersonally comparable.[107]

Arrow leaves open the possibility of relaxing any of the two aspects of condition *IIA*, which shows that the proposed solutions do not seem entirely satisfactory to him. We consider that the diverse alternatives in the literature of social choice theory that try to overcome Arrow's impossibility theorem are value-laden in one way or another. Arrow's approach leads economists (including himself) to make value judgments in order to find a way to obtain a "good" social choice procedure.

3.5 Conclusion

In this chapter we identified the three main points of entry that lead Arrow to develop his impossibility theorem: (1) the standard theory of the firm, (2) electoral politics, and (3) game theory. The three points of entry help us understand how Arrow fine-tuned his thesis topic until he got his impossibility theorem. The first point of entry is determined by the academic context

[107] According to Dasgupta (2021), Sen's treatment of interpersonal comparisons, which has been widely accepted in the literature as a way of enriching the informational basis of social choice, is only valid if we interpret a social welfare function as representing the social welfare judgment of a person, not as a voting rule. The following quote illustrates Dasgupta's position:
"Sen's principal accomplishment was to show what must be assumed in relation to the measurability and interpersonal comparability of individual well-being for an ethical system to be coherent. The reverse problem, of determining ethical thought, which is the counterpart of Arrow's question regarding voting rules, is harder" (Dasgupta 2021, 420).
Dasgupta's proposal is interesting because it shows that a person's ethical reasoning to discover her social preferences over alternative policies is not the same as the ethical reasoning to identify voting rules. However, we do not focus on studying the articulation between the former ("normative social choice") and the latter ("political social choice") in this chapter.

because Arrow shows his interest in creating a statistical model based on the work of John Hicks. The other two points of entry are more associated with the geopolitical context characterized by the beginning of the Cold War.

Although his fascination with mathematical logic and voting procedures were important for obtaining his impossibility result, we can say that the academic context (at the City College of New York, at Columbia University, at the Cowles Commission, and at the Rand Corporation) in which he developed his ideas also played a remarkable role. We can mention some facts that were in one way or another important in obtaining Arrow's impossibility theorem: the influence of Alfred Tarski's course in the development of Arrow's axiomatic approach, Harold Hotelling's support for obtaining a fellowship for his Ph.D. studies in economics, and Olaf Helmer's question about the application of game theory when players are nations and not individuals.

In this chapter we show that value judgments not only influence the choice of the topic and the goals of Arrow but also have an impact on the interpretation of his impossibility theorem. We corroborate the hypothesis proposed by Mongin (2006b) about Arrow's position with respect to value judgments: Arrow can be considered as a weak neutrality theorist. This not only means that Arrow can logically make value judgments, but also that *qua economist* he might and even should make value judgments. On the one hand, Arrow makes a logical analysis of the consistency of different value judgments in mechanisms of social choice. On the other hand, he makes value judgments by himself, not only to analyze each of the conditions of his theorem but also when relaxing one of the conditions in order to escape from the impossibility result.

The fact of studying Arrow's impossibility theorem under Mongin's prism allowed us, on the one hand, to examine to what extent Mongin's neutrality theses are useful to determine the value judgments made by the economist *qua economist*. On the other hand, it allowed us to consider that different ways of interpreting the role of value judgments in Arrow's impossibility theorem are possible and deserve to be studied.

General Conclusion

The aim of this dissertation was to determine the role played by value judgments in the study of normative economics. We have shown throughout the three chapters that economists cannot dispense with value judgments in the study of normative economics.

Review of the Chapters

In Chapter 1, we treated two questions: (1) To what extent are economists willing to make value judgments for themselves when studying normative economics? (2) How do economists take into account the value judgments of economic agents in the study of normative economics? We have shown that a theoretical framework is necessary to study the distinction between the value judgments of economists and those of economic agents in normative economics. For this reason, two approaches have been useful: Sen's classification of value judgments and Mongin's theses about value neutrality in economics. Making this distinction explicit allows us to determine, on the one hand, that economists cannot avoid making value judgments in their professional work, and on the other hand, that it is important to consider the normative reasons that lead economic agents to justify their preferences.

In Chapter 2, we chose the value of freedom among many important values with the aim of establishing how economists study this value and how economic agents use this value to make their choices. More precisely the question was: (3) How do normative economists study and characterize a value (freedom) in relation with the value judgments of economic agents (as represented by their preferences)? For this purpose, we discussed the importance of preference-based approaches in the measuring of freedom of choice. This led us to analyze how these approaches use the tools of social choice theory to develop the conception of freedom in economics. It resulted in the identification of three ways in which social choice theory influences the study of preference-based approaches. First, the importance of preferences in

social choice theory is essential to establish the connection between freedom and individual preferences. Second, the important place that consequences have in social choice theory leads to consider the process aspect of freedom and the opportunity aspect of freedom. Third, the notions of incompleteness and maximality that come from social choice theory play an important role in the preference rankings for the evaluation of freedom.

From the study of a single value, we turned in Chapter 3 to the study of the role of value judgments in the elaboration and interpretation of Arrow's impossibility theorem. More precisely, we addressed the following question: (4) To what extent are value judgments important in the elaboration and interpretation of normative economic theories? We have shown that value judgments do not only influence the choice of the topic and the goals of Arrow, but also have an impact on the interpretation of his impossibility theorem. On the one hand, we identified how Arrow's fascination with mathematical logic and voting procedures, and his academic context were important for obtaining the impossibility result. On the other hand, we showed that Mongin's value neutrality classification studied in the first chapter was useful to determine that Arrow could not dispense with value judgments when stating his impossibility theorem and its possible solutions.

Future directions

The present dissertation may be extended in different ways. First, an extension of this research is to study the role that value judgments play in the progress of normative economics. We take as a starting point the provisional definition of progress given by Mongin.

> A shift from theory T to theory T' is progressive if: (1) T' provides a solution to at least one unresolved problem of T; (2) T' provides a solution to the main problems that T had already addressed and resolved in its own way; (3) T' raises new problems and manages to solve at least one of them; (4) T does not satisfy the previous conditions with respect to T' (Mongin 2006a, 25)

Mongin's conception does not make explicit the role of value judgments of economists in the progress of normative economics. We think that it would be interesting to determine how the value judgments of economists influence these four requirements that allow us to know if there is progress from one form of normative economics to another form. This would allow us to

better understand how normative economics has progressed and how such progress can confirm its status as a science.

A second aspect that could be extended is to corroborate through a case study the importance of the distinction between the value judgments of economists and those of economic agents. We think that this distinction can play an essential role in the study of markets that can be considered repugnant. Roth (2007) affirms that the repugnance of transactions can shape a market and the choices people make. Among the transactions that can be considered inappropriate by people in some times and places we have the following: transactions to eat horse or dog meat, buying or selling human organs, buying a mother's child, dwarf tossing, various forms of price gouging, etc. We are interested in exploring the debate on the advantages and disadvantages of markets for human organs (Becker and Elias 2007; Elias, Lacetera, and Macis 2016; Roscoe 2013). It is important to analyze the normative arguments against the organ market such as: the commodification of the human body, the discouragement of altruistic donations, the coercion that the poorest can suffer to sell their organs, among others (Hansmann 1989; Thorne 2006). It is necessary to consider the different ways that have already been proposed by economists to regulate the organ market in order to be accepted by society. As a result, we can mention the control of organ prices by the State, the application of forms of non-monetary compensation to donors, the authorization by the State of persons or institutions who buy or sell organs, the restriction on the type of organ to be sold in the market, etc. (Cohen 2014).

A third aspect that could be extended is the study of an area that conceive freedom as a normative criterion in economics. We refer to the theory of rights introduced by Amartya Sen in his article "The Impossibility of a Paretian Liberal", published in 1970. This impossibility is based on the incompatibility of three conditions: unrestricted domain, the weak Pareto principle, and the condition of minimal liberty. Although there is an extensive literature devoted to either dropping at least one of these conditions or loosening at least one of them, we consider that it would be important to analyze the role of value judgments in Sen's elaboration and interpretation of the Paretian liberal paradox. This will allow us to understand how Sen's own values and the cultural context influenced his methodological choice, and the role of value judgments in the formulation of the conditions of his paradox.

A fourth aspect that could be extended is to study the articulation between "normative" social choice and "political" social choice. This would lead us to study the difference between a

person's ethical reasoning to discover her social preferences over alternative policies and her ethical reasoning to identify voting rules (Dasgupta 2021). The idea would be to study how individual's social preferences complement democratic voting rules to obtain a social decision. This will lead us to question whether the solution proposed by Sen (1970 [2017]) to escape from the impossibility theorem (considering interpersonal comparisons of utility to enrich the informational basis of social choice), is appropriate only in the case of studying the social welfare judgment of a person, and not in the case of a voting rule.

Appendix

We propose a proof of Arrow's theorem based on Sen (2011)[108]:

To do the proof, Sen starts by defining a set G of individuals which is "decisive" over the ordered pair of alternatives $\{x, y\}$ denoted $D_G(x, y)$. Then Sen (2011, 38–39) proposes the following two lemmas:

Lemma I: If $D_G(x, y)$ for any ordered pair of alternatives $\{x, y\}$, then D_G.

In order to establish this lemma, Sen shows that $D_G(x, y) \rightarrow D_G(c, d)$, for all c and d. He considers the case in which the states x, y, c, d are all distinct. He also assumes the following configuration of individual preferences: for all persons j in G: cP_jx, xP_jy, and yP_jd and for all persons i not in G: cP_ix, yP_id (nothing is assumed regarding the ranking of the other pairs). By $D_G(x, y)$, we get xPy, and by the weak Pareto Principle, we have cPx and yPd. In consequence, by transitivity, we obtain: cPd. By IIA, cPd must depend on individual preferences only over $\{c, d\}$, which implies $D_G(c, d)$.

Lemma II: For any G, if D_G and if G has more than one person in it and can be, thus, partitioned into two nonempty parts G_1 and G_2, then either D_{G1} or D_{G2}.

For all persons i in G_1: xP_iy, and xP_iz, with any feasible ranking of y, z, and that for all persons j in G_2: xP_jy, zP_jy, with any feasible ranking of x, z. Nothing is presumed from the preferences of persons not in G. Since every individual in G prefers x to y, then we must have xPy. If xPz, then G_1 members would be decisive over this pair since every individual in G_1 prefers x to z (every individual not in G_1 can rank this pair in any way). If G_1 is not to be decisive, we must have zRx for some set of individual preferences over x, z of persons not in G_1. But

[108] With respect to the formal proof, there are various versions, including Arrow (1963), Sen (1970 [2017], 2014),Campbell and Kelly (2002), Gaertner (2009), among others.

$zRx \,\&\, xPy \rightarrow zPy$. Since only G_2 members prefer z to y, this implies that G_2 is decisive over this pair. By Lemma I, G_2 is generally decisive. Then either G_1 or G_2 must be decisive.

By the weak Pareto Principle, any group is socially decisive over every pair of social states. By partitioning into two nonempty parts, and each time choosing the decisive part, we attain a decisive individual, who must be a dictator.

Summary

The aim of this dissertation is to determine the role played by value judgments in the study of normative economics. We show throughout three chapters that economists cannot dispense with value judgments in the study of normative economics.

In Chapter 1, we treat two questions: (1) To what extent are economists willing to make value judgments for themselves when studying normative economics? (2) How do economists take into account the value judgments of economic agents in the study of normative economics? We suggest that a theoretical framework is necessary to study the distinction between the value judgments of economists and those of economic agents in normative economics. Making this distinction explicit allows us to determine, on the one hand, that economists cannot avoid making value judgments in their professional work, and on the other hand, that it is important to consider the normative reasons that lead economic agents to justify their preferences.

In Chapter 2, we treat the following question: (3) How do normative economists study and characterize a value (freedom) in relation with the value judgments of economic agents (as represented by their preferences)? For this purpose, we discuss the importance of preference-based approaches in the measuring of freedom of choice. This leads us to analyze how these approaches use the tools of social choice theory to develop the conception of freedom in economics. We suggest that the importance of preferences, consequences, and notions such as incompleteness and maximality in social choice theory play an important role in the study of preference-based approaches.

From the study of a single value, we turn in Chapter 3 to the study of the role of value judgments in the elaboration and interpretation of Arrow's impossibility theorem. We address the following question: (4) To what extent are value judgments important in the elaboration and interpretation of normative economic theories? We identify that Arrow's fascination with mathematical logic and voting procedures, and his academic context were important for obtaining his impossibility result. On the other hand, we suggest that Mongin's value neutrality classification studied in the first chapter is useful to determine that Arrow cannot dispense with value judgments when stating his impossibility theorem and its possible solutions.

Keywords: *freedom of choice − interpersonal comparisons of utility − economic agent − economist − normative economics − preference − social choice theory − value judgment − value neutrality*

JEL codes: A13, B41, D60, D70

Résumé

L'objectif de cette thèse est de déterminer le rôle que jouent les jugements de valeur dans l'étude de l'économie normative. Au long des trois chapitres de ce travail, nous montrons que les économistes ne peuvent pas se passer de jugements de valeur dans l'étude de l'économie normative.

Dans le chapitre 1, nous abordons deux questions : (1) Dans quelle mesure les économistes assument-ils de porter des jugements de valeur lorsqu'ils étudient l'économie normative ? (2) Comment les économistes prennent-ils en compte les jugements de valeur des agents économiques dans l'étude de l'économie normative ? Nous défendons l'idée qu'un cadre théorique est nécessaire pour étudier la distinction entre les jugements de valeur des économistes et ceux des agents économiques en économie normative. Rendre explicite cette distinction permet de constater, d'une part, que les économistes ne peuvent éviter de porter des jugements de valeur dans leur activité professionnelle, et d'autre part, qu'il est important de considérer les raisons normatives qui conduisent les agents économiques à justifier leurs préférences.

Dans le chapitre 2, nous abordons la question suivante : (3) Comment les économistes normatifs étudient-ils et caractérisent-ils une valeur (la liberté) en relation avec les jugements de valeur des agents économiques (représentés par leurs préférences) ? Pour ce faire, nous discutons l'importance des approches visant à mesurer la liberté de choix qui se basent sur les préférences des agents. Cela nous amène à analyser comment ces approches utilisent les outils de la théorie du choix social pour développer la conception de la liberté. Nous défendons l'idée que l'importance des préférences, des conséquences et des notions telles que l'incomplétude et la maximalité dans la théorie du choix social joue un rôle important dans l'étude des approches basées sur les préférences.

Nous élargissons notre analyse en passant, dans le chapitre 3, de l'étude de la seule valeur de liberté au rôle des jugements de valeur en général en nous focalisant sur le théorème d'impossibilité d'Arrow. Nous abordons la question suivante : (4) Dans quelle mesure les jugements de valeur sont-ils importants dans l'élaboration et l'interprétation des théories économiques normatives ? Nous mettons en évidence le fait que la fascination d'Arrow pour la logique mathématique, les procédures de vote et le contexte académique ont été importantes pour obtenir son résultat d'impossibilité. D'autre part, nous suggérons que la classification de neutralité des valeurs de Mongin étudiée dans le premier chapitre de cette thèse est utile pour déterminer qu'Arrow n'est pas libre de tout jugement de valeur lorsqu'il énonce son théorème d'impossibilité et ses solutions possibles.

Mots clés : liberté de choix −comparaisons interpersonnelles d'utilité −agent économique − économiste − économie normative − préférence − théorie du choix social − jugement de valeur − neutralité axiologique

Bibliography

Alvarez, Maria. 2017. *Reasons for Action: Justification, Motivation, Explanation*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Winter ed. https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/.

Amadae, Sonja. 2003. *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism*. Chicago: The University of Chicago Press.

Arneson, Richard. 1998. "Real Freedom and Distributive Justice." In *Freedom in Economics. New Perspectives in Normative Analysis*, edited by Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy, 165–96. London: Routledge.

Arrow, Kenneth J. 1963. *Social Choice and Individual Values*. 3rd ed. New Haven: Yale University.

———. 1973. "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice." *Journal of Philosophy* 70 (9): 245–63.

———. 1977a. "Current Developments in the Theory of Social Choice." *Rationality, Choice, and Morality* 44 (4): 607–22.

———. 1977b. "Extended Sympathy and the Possibility of Social Choice." *American Economic Review* 67 (1): 219–25.

———. 1984a. "A Difficulty in the Concept of Social Welfare." In *Collected Papers of Kenneth J. Arrow. Social Choice and Justice*, 1–29. Oxford, England: Basil Blackwell.

———. 1984b. "The Principle of Rationality in Collective Decisions." In *Collected Papers of Kenneth J. Arrow. Social Choice and Justice*, 45–58. Oxford, England: Basil Blackwell.

———. 1984c. "Values and Collective Decision Making." In *Collected Papers of Kenneth J. Arrow. Social Choice and Justice*, 59–80. Oxford, England: Basil Blackwell.

———. 1995. "A Note on Freedom and Flexibility." In *Choice, Welfare and Development. A Festschrift in Honour of Amartya K. Sen*, edited by Kaushik Basu, Prasanta K. Pattanaik, and Kotaro Suzumura, 7–16. Oxford: Clarendon Press.

———. 2014a. "Commentary." In *The Arrow Imposibility Theorem*, edited by Eric Maskin and Amartya Sen, 57–65. New York: Columbia University Press.

———. 2014b. "The Origins of The Impossibility Theorem." In *The Arrow Imposibility Theorem*, edited by Eric Maskin and Amartya Sen, 143–48. New York: Columbia University Press.

Arrow, Kenneth J., Kristen Monroe, and Nicholas Lampros. 2017. *On Ethics And Economics. Conversations with Kenneth J. Arrow*. New York: Routledge.

Arrow, Kenneth J., Amartya Sen, and Kotaro Suzumura. 2011. “Kenneth Arrow on Social Choice Theory.” In *Handbook of Social Choice and Welfare*, edited by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, 2:3–27. Amsterdam: North-Holland. https://doi.org/10.1016/S0169-7218(10)00013-4.

Backhouse, Roger E. 2005. “Economists, Values and Ideology: A Neglected Agenda.” *Revue de Philosophie Économique* 11: 31–55.

———. 2021. “Individualism and Ethics: Samuelson’s Welfare Economics.” In *Welfare Theory, Public Action, and Ethical Values: Revisiting the History of Welfare Economics*, edited by Roger E Backhouse, Antoinette Baujard, and Tamotsu Nishizawa, 186–207. Cambridge: Cambridge University Press.

Backhouse, Roger E, Antoinette Baujard, and Tamotsu Nishizawa. 2021. “Introduction: Revisiting the History of Welfare Economics.” In *Welfare Theory, Public Action, and Ethical Values: Revisiting the History of Welfare Economics*, edited by Roger E Backhouse, Antoinette Baujard, and Tamotsu Nishizawa, 1–18. Cambridge: Cambridge University Press.

Badiei, Sina. 2022. *Économie positive et économie normative chez Marx, Mises, Friedman et Popper*. Editions Matériologiques.

Badiei, Sina, Gilles Campagnolo, and Agnès Grivaux. 2022. “Introduction.” In *Le positif et le normatif en philosophie économique*, edited by Sina Badiei, Gilles Campagnolo, and Agnès Grivaux, 7–38. Paris: Editions Matériologiques.

Balinski, Michel, and Rida Laraki. 2007. “A Theory of Measuring, Electing, and Ranking.” *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 104 (21): 8720–25.

Barberà, Salvador, and Birgit Grodal. 2011. “Preference for Flexibility and the Opportunities of Choice.” *Journal of Mathematical Economics* 47 (3): 272–78. https://doi.org/10.1016/j.jmateco.2010.12.009.

Baujard, Antoinette. 2017. “L’économie du bien-être est morte . Vive l’économie du bien - être !” In *Philosophie économique. Un état de lieux.*, edited by Gilles Campagnolo and Jean-Sébastien Gharbi, 77–128. Paris: Editions Matériologiques.

Bavetta, Sebastiano. 2004. “Measuring Freedom of Choice : An Alternative View of a Recent Literature.” *Social Choice and Welfare* 22 (1): 29–48. https://doi.org/10.1007/s00355-003-0275-1.

Becker, Gary S, and Julio Jorge Elias. 2007. “Introducing Incentives in the Market for Live and Cadaveric Organ Donations.” *Journal of Economic Perspectives* 21 (3): 3–24.

Bergson, Abram. 1938. “A Reformulation of Certain Aspects of Welfare Economics.” *The Quarterly Journal of Economics* 52 (2): 310–34. https://doi.org/10.2307/1881737.

———. 1954. “On the Concept of Social Welfare.” *The Quarterly Journal of Economics* 68 (2): 233–52.

Berlin, Isaiah. 2002. *Liberty*. Edited by Henry Hardy. 2nd ed. Oxford: Oxford University Press. https://doi.org/10.4324/9781315091822-3.

Berthonnet, Irène, and Vincent Desreumaux. 2014. “La Coordination par le marché dans Social Choice and Individual Values : mérites normatifs ou mérites techniques ?” *Cahiers d’économie politique* 66 (1): 95–126. https://doi.org/10.3917/cep.066.0095.

Binder, Constanze. 2019. *Agency , Freedom and Choice*. Rotterdam: Springer.

Blaug, Mark. 1992. “The Distinction between Positive and Normative Economics.” In *The Methodology of Economics or, How Economists Explain*, 2nd ed., 112–34. Cambridge: Cambridge University Press. https://doi.org/10.3917/redp.282.0191.

———. 1998. “The Positive-Normative Distinction.” In *Handbook of Economic Methodology*, edited by John B Davis, Wade Hands, and Uskali Mäki, 370–74. Cheltenham, the UK: Edward Elgar.

Bordes, Georges, and Maurice Salles. 1978. “Sur l’impossibilité des fonctions de décision collective : un commentaire et un résultat.” *Revue d’économie politique* 88 (3): 442–48.

Boumans, Marcel, and John B Davis. 2010. “Value Judgments in Economics.” In *Economic Methodology Understanding Economics as a Science*, 169–94. New York: Palgrave Macmillan.

Broome, John. 1993. “A Cause of Preference Is Not an Object of Preference.” *Social Choice and Welfare* 10 (2): 57–68.

Campbell, Donald E., and Jerry S. Kelly. 2002. “Impossibility Theorems in The Arrovian Framework.” In *Handbook of Social Choice and Welfare*, edited by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, 1:36–94. Amsterdam: North-Holland.

Carter, Ian. 1995. “The Independent Value of Freedom.” *Ethics* 105 (4): 819–45. https://doi.org/10.1086/293754.

———. 1996. “The Concept of Freedom in The Work of Amartya Sen: An Alternative Analysis Consistent With Freedom’s Independent Value.” *Notizie Di Politeia* 43 (4): 7–22.

———. 1999. *A Measure of Freedom*. Oxford: Oxford University Press.

———. 2004. “Choice, Freedom, and Freedom of Choice.” *Social Choice and Welfare* 22 (1): 61–81. https://doi.org/10.1007/s00355-003-0277-z.

———. 2019. *Positive and Negative Liberty*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Winter ed. https://plato.stanford.edu/archives/win2019/entries/liberty-positive-negative/.

Caruso, Gregg. 2018. *Skepticism about Moral Responsibility*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Spring ed. https://doi.org/10.1111/papq.12197.

Christman, John. 1991. “Liberalism and Individual Positive Freedom.” *Ethics* 101 (2): 343–59. https://doi.org/10.4324/9781315625546.

Cohen, I. Glenn. 2014. “Regulating the Organ Market: Normative Foundations for Market Regulation.” *Law and Contemporary Problems* 77 (3): 71–100.

Colander, David, and Huei Chun Su. 2015. “Making Sense of Economists’ Positive-Normative Distinction.” *Journal of Economic Methodology* 22 (2): 157–70. https://doi.org/10.1080/1350178X.2015.1024877.

Dasgupta, Partha. 2021. “The Perils of Cosmopolitan Intellectualism.” *Society* 58 (5): 416–23. https://doi.org/10.1007/s12115-021-00640-5.

Davis, John B. 2013. “Economists’ Odd Stand on the Positive-Normative Distinction: A Behavioral Economics View.” In *Handbook on Professional Economic Ethics: View from the Economics Profession and Beyond*, edited by George DeMartino and Deirdre McCloskey, 200–218. Oxford University Press. https://doi.org/http://dx.doi.org/10.2139/ssrn.2230062.

Dietrich, Franz, and Christian List. 2013. “A Reason-Based Theory of Rational Choice.” *Noûs* 47 (1): 104–34. https://doi.org/10.1111/j.1468-0068.2011.00840.

Dowding, Keith, and Martin van Hees. 2009. “Freedom of Choice.” In *Handbook of Rational and Social Choice*, edited by Paul Anand, Prasanta K. Pattanaik, and Clemens Puppe, 374–92. Oxford: Oxford University Press.

Elias, Julio, Nicola Lacetera, and Mario Macis. 2016. “Efficiency-Morality Trade-Offs in Repugnant Transactions: A Choice Experiment.”

Elster, Jon, and John Roemer. 1991. *Interpersonal Comparisons of Well-Being*. Cambridge: Cambridge University Press.

Escoffier, Bruno, Jérôme Lang, and Meltem Öztürk. 2008. “Single-Peaked Consistency and Its Complexity.” In *18th European Conference on Artificial Intelligence (ECAI’08)*, 1–14.

Feiwel, George R. 1987a. *Arrow and the Ascent of Modern Economic Theory*. London: Macmillan.

———. 1987b. *Arrow and the Foundations of The Theory of Economic Policy*. London: Macmillan.

Fisher, Donald. 1913. “The Problem of the Value-Judgment.” *The Philosophical Review* 22 (6): 623–38.

Fleurbaey, Marc. 2003. “On the Informational Basis of Social Choice.” *Social Choice and Welfare* 21 (2): 347–84. https://doi.org/10.1007/s00355-003-0263-5.

Fleurbaey, Marc, and Philippe Mongin. 2005. *The News of the Death of Welfare Economics Is Greatly Exaggerated*. *Social Choice and Welfare*. Vol. 25. https://doi.org/10.1007/s00355-005-0010-1.

Foster, James. 1993. “Notes on Effective Freedom.” *Oxford Poverty and Human Development Initiative*, no. 37: 10.

Frankfurt, Harry. 1971. “Freedom of the Will and the Concept of a Person.” *Journal of Philosophy* 68 (1): 5–20.

Friedman, Milton. 1953. “The Methodology of Positive Economics.” In *Essays in Positive Economics*, 3–43. Chicago: University of Chicago Press. https://doi.org/10.1017/CBO9780511581427.002.

Gaertner, Wulf. 2009. *A Primer in Social Choice Theory*. London: LSE.

Gravel, Nicolas, Jean-François Laslier, and Alain Trannoy. 1998. “Individual Freedom of Choice in a Social Setting.” In *Freedom in Economics. New Perspectives in Normative Analysis*, edited by Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy, 76–92. London: Routledge.

Grünberg, Ludwig. 1978. “Rationality and The Basis of The Value Judgment.” *The Journal of Value Inquiry* 12 (2): 126–33. https://doi.org/10.1007/BF00145889.

Gul, Faruk, and Wolfgang Pesendorfer. 2008. “The Case for Mindless Economics.” In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin and Andrew Schotter, 3–39. New York: Oxford University Press.

Hamlin, Aaron. 2012. “Interview with Nobel Laureate Dr. Kenneth Arrow.” The Center for Election Science. 2012. https://electionscience.org/commentary-analysis/voting-theory-podcast-2012-10-06-interview-with-nobel-laureate-dr-kenneth-arrow/.

Hammond, Peter J. 1977. “Dual Interpersonal Comparisons of Utility and The Welfare Economics of Income Distribution.” *Journal of Public Economics* 7: 51–71.

———. 1987. “Social Choice: The Science of the Impossible?” In *Arrow and the Foundations of the Theory of Economic Policy*, edited by George R. Feiwel, 116–31. London: Macmillan.

———. 1991. “Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made.” In *Interpersonal Comparisons of Well-Being*, edited by Jon Elster and John Roemer, 200–254. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139172387.008.

Hands, Wade. 2012. “The Positive-Normative Dichotomy and Economics.” In *Handbook of the Philosophy of Science*, edited by Uskali Mäki, 13:219–39.

Hansmann, H. 1989. “The Economics and Ethics of Markets for Human Organs.” *Journal of Health Politics, Policy and Law* 14 (1): 57–85. https://doi.org/10.1215/03616878-14-1-57.

Hare, R M. 1952. *The Language of Morals*. Oxford: Clarendon Press.

Harsanyi, John. 1955. “Cardinal Welfare , Individualistic Ethics , and Interpersonal Comparisons of Utility.” *Journal of Political Economy* 63 (4): 309–21.

———. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

Hausman, Daniel M. 2012. *Preference, Value, Choice, and Welfare*. New York: Cambridge University Press.

Hibbard, Jane. 2011. “Kenneth Joseph Arrow. Oral History Interviews.” https://doi.org/10.33043/th.1.2.65-69.

Hicks, John. 1940. “The Evaluation of the Social Income.” *Economica* 7: 105–24.

Hoefer, Carl. 2016. *Causal Determinism*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Spring ed. https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/.

Hume, David. 1960. *A Treatise on Human Nature*. Edited by Lewis Amherst Selby-Bigge. Oxford: Clarendon Press.

Hutchison, T W. 1964. *“Positive” Economics and Policy Objectives*. London: Routledge. Taylor & Francis Group.

Igersheim, Herrade. 2017. “The Death of Welfare Economics : History of a Controversy.” *The Center for the History of Political Economy Working Paper Series*, 1–49. https://doi.org/http://dx.doi.org/10.2139/ssrn.2901574.

Jones, Peter, and Robert Sugden. 1982. “Evaluating Choices.” *International Review of Law and Economics* 2 (1): 47–65.

Kaldor, Nicholas. 1939. “Welfare Propositions in Economics and Interpersonal Comparisons of Utility.” *Economic Journal* 49: 549–52.

Kaneko, M. 1984. “Social Choice with Interpersonal Utility Comparisons.” *Social Choice and Welfare*, no. 1: 165–75.

Keynes, John Neville. 1917. *The Scope and Method of Political Economy*. *The Scope and Method of Political Economy*. 4th ed. London: Macmillan. https://doi.org/10.4324/9781315134802.

Klemisch-Ahlert, Marlies. 1993. “Freedom of Choice: A Comparison of Different Rankings of Opportunity Sets.” *Social Choice and Welfare* 10 (3): 189–207.

Kolm, Serge-Christophe. 1994. “The Meaning of ‘Fundamental Preferences.’” *Social Choice and Welfare* 11 (3): 193–98.

Kornai, János. 1988. “Individual Freedom and Reform of the Socialist Economy.” *European Economic Review* 32 (2–3): 233–67. https://doi.org/10.1016/0014-2921(88)90172-9.

Kreps, David. 1979. “A Representation Theorem for " Preference for Flexibility ".” *Econometrica* 47 (3): 565–77.

Laslier, Jean-François, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. 1998. *Freedom in Economics. New Perspective in Normative Analysis*. Edited by Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. London: Routledge.

Little, Ian. 1952. “Social Choice and Individual Values.” *Journal of Political Economy* 60 (5): 422–32.

MacCallum, Gerald. 1967. “Negative and Positive Freedom.” *The Philosophical Review* 76 (3): 312–34.

MacKay, Alfred. 1986. “Extended Sympathy and Interpersonal Utility Comparisons.” *Journal of Philosophy* 83 (6): 305–22.

Małecka, Magdalena. 2021. “Values in Economics: A Recent Revival with a Twist.” *Journal of Economic Methodology* 28 (1): 88–97. https://doi.org/10.1080/1350178X.2020.1868776.

Maskin, Eric, and Amartya Sen. 2014. *The Arrow Impossibility Theorem*. New York: Columbia University Press.

Mill, John Stuart. 1967. “On the Definition of Political Economy; and on the Method of Investigation Proper to It.” In *Collected Works of John Stuart Mill*, edited by J. M. Robson, 4:276–97. Toronto: Toronto University Press. https://doi.org/10.2307/2343853.

———. 1998. *On Liberty*. Edited by Jim Manis. Pennsylvania: The Pennsylvania State University.

Mongin, Philippe. 1999. “Normes et jugements de valeur en économie normative.” *Information sur les Sciences Sociales* 38 (4): 521–53. https://doi.org/0803973233.

———. 2001a. “The Impartial Observer Theorem of Social Ethics.” *Economics and Philosophy* 17 (2): 147–79. https://doi.org/10.1017/S0266267101000219.

———. 2001b. “Value Judgments and Value Neutrality in Economics. A Perspective from Today.” Paris. https://doi.org/10.1111/j.1468-0335.2006.00501.x.

———. 2002. “Is There Progress in Normative Economics?” In *Is There Progress in Economics? Knowledge, Truth and History of Economic Thought*, edited by Stephan Boehm, Christian Gehrke, Heinz Kurz, and Richard Sturn, 145–72. Cornwall: Edward Elgar Publishing.

———. 2006a. “A Concept of Progress for Normative Economics.” *Economics and Philosophy* 22 (1): 19–54. https://doi.org/10.1017/S0266267105000696.

———. 2006b. “Value Judgments and Value Neutrality in Economics.” *Economica* 73 (290): 257–86. https://doi.org/10.1111/j.1468-0335.2006.00501.x.

———. 2018. “Les origines de la distinction entre positif et normatif en économie.” https://doi.org/http://dx.doi.org/10.2139/ssrn.3106604.

Morgan, Mary. 2012. *The World in the Model: How Economists Work and Think*. *The World in the Model: How Economists Work and Think*. https://doi.org/10.1017/CBO9781139026185.

Morreau, Michael. 2019. *Arrow’s Theorem*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Winter ed. https://plato.stanford.edu/archives/win2019/entries/arrows-theorem/.

Myrdal, Gunnar. 1958. *Value in Social Theory*. London: Routledge.

Nehring, Klaus, and Clemens Puppe. 2002. “A Theory of Diversity.” *Econometrica* 70 (3): 1155–98.

———. 2009. “Diversity.” In *The Handbook of Rational & Social Choice*, edited by Paul Anand, Prasanta K. Pattanaik, and Clemens Puppe, 298–322. Oxford: Oxford University Press.

O’Connor, Timothy, and Christopher Franklin. 2021. *Free Will*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Spring ed. https://plato.stanford.edu/archives/spr2021/entries/freewill/.

Pattanaik, Prasanta K. 2014. “Introduction.” In *The Arrow Imposibility Theorem*, edited by Eric Maskin and Amartya Sen, 1–21. New York: Columbia University Press.

Pattanaik, Prasanta K., and Yongsheng Xu. 1990. “On Ranking Opportunity Sets in Terms of Freedom of Choice.” *Louvain Economic Review* 56 (3/4): 383–90.

———. 1998. “On Preference and Freedom.” *Theory and Decision* 44 (2): 173–98. https://doi.org/10.1023/A:1004924211553.

———. 2000. “On Diversity and Freedom of Choice.” *Mathematical Social Sciences* 40: 123–30.

Pazner, Elisha. 1979. “Equity, Nonfeasible Alternatives and Social Choice: A Reconsideration of the Concept of Social Welfare.” In *Aggregation and Revelation of Preferences*, edited by Jean-Jacques Laffont. Amsterdam: North-Holland.

Pettit, Philip. 1997. *Republicanism: A Theory of Freedom and Government*. Edited by David Miller and Alan Ryan. Oxford: Oxford University Press.

Pigou, Arthur. 1920. *The Economics of Welfare*. London: Macmillan.

Puppe, Clemens. 1996. “An Axiomatic Approach to ‘Preference for Freedom of Choice.’” *Journal of Economic Theory* 68 (1): 174–99. https://doi.org/10.1006/jeth.1996.0009.

Putnam, Hilary. 2002. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, the US: Harvard University Press.

Putnam, Hilary, and Vivian Walsh. 2012. *The End of Value-Free Economics*. London: Routledge. Taylor & Francis Group. https://doi.org/10.4324/9780203154007.

Rawls, John. 1981. “The Basic Liberties and Their Priority.” In *The Tanner Lecture on Human Values,* 3:3–87. Salt Like City: University of Utah Press.

———. 1999. *A Theory of Justice*. Revised ed. Cambridge: The Belknap press.

Reiss, Julian. 2017. “Fact-Value Entanglement in Positive Economics.” *Journal of Economic Methodology* 24 (2): 134_149. https://doi.org/10.1080/1350178X.2017.1309749.

Robbins, Lionel. 1984. *An Essay on the Nature and Significance of Economic Science*. 3rd ed. London: Macmillan.

Röpke, Wilhelm. 2015. “A Value Judgment on Value Judgments.” *Journal of Markets & Morality* 18 (2): 497–514.

Roscoe, Philip. 2013. “On the Possibility of Organ Markets and the Performativity of Economics.” *Journal of Cultural Economy* 6 (4): 386–401. https://doi.org/10.1080/17530350.2013.772069.

Rosenbaum, Eckehard F. 2000. “On Measuring Freedom.” *Journal of Theoretical Politics* 12 (2): 205–27.

Roth, Alvin E. 2007. “Repugnance as a Constraint on Markets.” *Journal of Economic Perspectives* 21 (3): 37–58. https://doi.org/10.1257/jep.21.3.37.

Saito, Nao. 2011. “Arrow’s Social Ordering as Moral Obligation.” *G-COE GLOPE II Working Paper Series Working Paper N°51*.

Salles, Maurice. 1976. “Sur l’impossibilité des fonctions de décision collective.” *Revue d’économie politique* 86 (2): 269–76.

Samuelson, Paul A. 1947. *Foundantions of Economic Analysis*. MA: Harvard University Press.

———. 1981. “Bergsonian Welfare Economics.” In *Economic Welfare and The Economics of Soviet Socialism: Essays in Honor of Abram Bergson,* edited by Steven Rosefielde, 223–66. Cambridge: Cambridge University Press.

Scanlon, Thomas. 1991. “The Moral Basis of Interpersonal Comparisons.” In *Interpersonal Comparisons of Well-Being*, edited by Jon Elster and John Roemer, 17–44. Cambridge: Cambridge University Press.

Scarantino, Andrea. 2009. “On the Role of Values in Economic Science: Robbins and His Critics.” *Journal of the History of Economic Thought* 31 (4): 449–73. https://doi.org/10.1017/S1053837209990253.

Scitovsky, Tibor. 1941. “A Note on Welfare Propositions in Economics.” *Review of Economic Studies* 9: 77–88.

Sen, Amartya. 1967. “The Nature and Classes of Prescriptive Judgments.” *The Philosophical Quarterly* 17 (66): 46–62.

———. 1969. “Quasi-Transitivity, Rational Choice and Collective Decisions.” *The Review of Economic Studies* 36 (3): 381–93.

———. 1970. “The Impossibility of a Paretian Liberal.” *Journal of Political Economy* 78 (1): 152–57. https://doi.org/10.1086/260083.

———. 1976. “Liberty , Unanimity and Rights.” *Economica* 43 (171): 217–45.

———. 1977a. “Rational Fools : A Critique of the Behavioral Foundations of Economic Theory.” *Philosophy and Public Affairs* 6 (4): 317–44.

———. 1977b. “Social Choice Theory : A Re-Examination.” *Econometrica* 45 (1): 53–88.

———. 1983. “Liberty and Social Choice.” *Journal of Philosophy* 80 (1): 5–28. https://doi.org/10.2307/3655750.

———. 1985. “Well-Being, Agency and Freedom: The Dewey Lectures 1984.” *Journal of Philosophy* 82 (4): 169–221.

———. 1988. “Freedom of Choice: Concept and Content.” *European Economic Review* 32 (2–3): 269–94.

———. 1990. “Welfare, Freedom and Social Choice: A Reply.” *Louvain Economic Review* 56 (3/4): 451–85.

———. 1991. “Welfare , Preference and Freedom.” *Journal of Econometrics* 50 (1–2): 15–29.

———. 1993a. “Internal Consistency of Choice.” *Econometrica* 61 (3): 495–521.

———. 1993b. “Markets and Freedoms : Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms.” *Oxford Economic Papers* 45 (4): 519–41.

———. 1996. “Freedom, Capabilities and Public Action: A Response.” *Notizie Di Politeia* 43 (4): 107–25.

———. 1997a. “Individual Preference as the Basis of Social Choice.” In *Social Choice Re-Examined*, edited by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, 15–37. Palgrave Macmillan.

———. 1997b. “Maximization and The Act of Choice.” In *Rationality and Freedom*, edited by Amartya Sen, 158–205. Massachusetts: Harvard University Press.

———. 1997c. *On Economic Inequality*. Expanded E. Oxford: Clarendon Press.

———. 1999. “The Possibility of Social Choice.” *American Economic Review* 89 (3): 349–78.

———. 2002. *Rationality and Freedom*. Massachusetts: Harvard University Press.

———. 2006. *Identity and Violence: The Illusion of Destiny*. New York: Norton & Company.

———. 2011. “The Informational Basis of Social Choice.” In *Handbook of Social Choice and Welfare*, edited by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, 2:29–46. Amsterdam: North-Holland. https://doi.org/10.1016/S0169-7218(10)00014-6.

———. 2014. “Arrow and The Impossibility Theorem.” In *The Arrow Imposibility Theorem*, edited by Eric Maskin and Amartya Sen, 29–42. New York: Columbia University Press.

———. 2017. *Collective Choice and Social Welfare*. Expanded e. London: Penguin Books.

Senior, Nassau. 1966. “Four Introductory Lectures on Political Economy.” In *Selected Writings on Economics by Nassau W. Senior, Reprints of Economic Classics*. New York: Augustus Kelley Publishers.

Skinner, Quentin. 2006. “A Third Concept of Liberty.” In *The Liberty Reader*, edited by David Miller, 243–54. New York: Routledge. Taylor & Francis Group.

Sugden, Robert. 1998. “The Metric of Opportunity.” *Economics and Philosophy* 14 (2): 307–37. https://doi.org/10.1017/S0266267100003874.

———. 2003. “Opportunity as a Space for Individuality: Its Value and the Impossibility of Measuring It.” *Ethics* 113 (4): 783–809.

———. 2010. “Opportunity as Mutual Advantage.” *Economics and Philosophy* 26 (1): 47–68. https://doi.org/10.1017/S0266267110000052.

Suppes, Patrick. 2005. “The Pre-History of Kenneth Arrow’s Social Choice and Individual Values.” *Social Choice and Welfare* 25 (2–3): 319–26. https://doi.org/10.1007/s00355-005-0006-x.

Suzumura, Kotaro. 2002. “Introduction.” Edited by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura. *Handbook of Social Choice and Welfare* 1: 1–32.

———. 2021. “John Hicks’s Farewell to Economic Welfarism: How Deeply Rooted and Far Reaching Is His Non-Welfarist Manifesto?” In *Welfare Theory, Public Action, and Ethical Values: Revisiting the History of Welfare Economics*, edited by Roger E Backhouse, Antoinette Baujard, and Tamotsu Nishizawa, 159–85. Cambridge: Cambridge University Press.

Thorne, Emanuel D. 2006. “The Economics of Organ Transplantation.” In *Handbook of the Economics of Giving, Altruism and Reciprocity*, edited by Serge-Christophe Kolm and Jean Mercier, 2:1335–70. https://doi.org/10.1016/S1574-0714(06)02020-3.

Unicef. 2022. “Female Genital Mutilation.” 2022. https://www.unicef.org/protection/female-genital-mutilation.

Van Hees, Martin. 2000. *Legal Reductionism and Freedom*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

———. 2004. “Freedom of Choice and Diversity of Options: Some Difficulties.” *Social Choice and Welfare* 22 (1): 253–66. https://doi.org/10.1007/s00355-003-0285-z.

Van Hees, Martin, and Marcel Wissenburg. 1999. “Freedom and Opportunity.” *Political Studies* 47 (1): 67–82. https://doi.org/10.1111/1467-9248.00188.

Von Neumann, John, and Oskar Morgenstern. 1953. *Theory of Games and Economic Behavior*. 3rd ed. London: Princeton University Press.

Weber, Max. 1949. *The Methodology of the Social Sciences*. Edited by Edward Shils and Henry Finch. Glencoe, Illinois: The Free Press. https://doi.org/10.2307/3707961.

———. 1965. *Le savant et le politique*. Paris: Plon.

Weitzman, Martin. 1992. "On Diversity." *Quartely Journal of Economics* 107 (2): 363–405.

Zimmerman, Michael, and Ben Bradley. 2019. *Intrinsic vs. Extrinsic Value*. Edited by Edward Zalta. *The Stanford Encyclopedia of Philosophy*. Spring ed. https://plato.stanford.edu/archives/spr2019/entries/value-intrinsic-extrinsic/.

Une étude du rôle des jugements de valeur en économie normative

L'objectif de cette thèse est de déterminer le rôle que jouent les jugements de valeur dans l'étude de l'économie normative. Au long des trois chapitres de ce travail, nous montrons que les économistes ne peuvent pas se passer de jugements de valeur dans l'étude de l'économie normative. Dans le chapitre 1, nous abordons deux questions : (1) Dans quelle mesure les économistes assument-ils de porter des jugements de valeur lorsqu'ils étudient l'économie normative ? (2) Comment les économistes prennent-ils en compte les jugements de valeur des agents économiques dans l'étude de l'économie normative ? Nous défendons l'idée qu'un cadre théorique est nécessaire pour étudier la distinction entre les jugements de valeur des économistes et ceux des agents économiques en économie normative. Dans le chapitre 2, nous abordons la question suivante : (3) Comment les économistes normatifs étudient-ils et caractérisent-ils une valeur (la liberté) en relation avec les jugements de valeur des agents économiques (représentés par leurs préférences) ? Pour ce faire, nous analysons comment ces approches visant à mesurer la liberté de choix sur la base des préférences individuelles utilisent les outils de la théorie du choix social pour développer la conception de la liberté en économie. Dans le chapitre 3, nous abordons la question suivante : (4) Dans quelle mesure les jugements de valeur sont-ils importants dans l'élaboration et l'interprétation des théories économiques normatives ? Nous mettons en évidence le fait qu'Arrow n'est pas libre de tout jugement de valeur lorsqu'il énonce son théorème d'impossibilité et ses solutions possibles.

Mots-clés en français : liberté de choix –comparaisons interpersonnelles d'utilité –agent économique –économiste – économie normative – préférence – théorie du choix social – jugement de valeur – neutralité axiologique

Elucidating the Role of Value Judgments in Normative Economics

The aim of this dissertation is to determine the role played by value judgments in the study of normative economics. We show throughout three chapters that economists cannot dispense with value judgments in the study of normative economics. In Chapter 1, we treat two questions: (1) To what extent are economists willing to make value judgments for themselves when studying normative economics? (2) How do economists take into account the value judgments of economic agents in the study of normative economics? We suggest that a theoretical framework is necessary to study the distinction between the value judgments of economists and those of economic agents in normative economics. In Chapter 2, we treat the following question: (3) How do normative economists study and characterize a value (freedom) in relation with the value judgments of economic agents (as represented by their preferences)? For this purpose, we analyze how preference-based approaches in the measuring of freedom of choice use the tools of social choice theory to develop the conception of freedom in economics. In Chapter 3, we study the role of value judgments in the elaboration and interpretation of Arrow's impossibility theorem. We address the following question: (4) To what extent are value judgments important in the elaboration and interpretation of normative economic theories? We suggest that Arrow cannot dispense with value judgments when stating his impossibility theorem and its possible solutions.

Mots-clés en anglais : *freedom of choice –interpersonal comparisons of utility –economic agent – economist – normative economics – preference – social choice theory – value judgment – value neutrality*

Discipline : SCIENCES ÉCONOMIQUES

Université de Reims Champagne-Ardenne

REGARDS - EA 6292

57 rue Pierre Taittinger - 51100 REIMS