How mathematical impossibility changed welfare economics
A history of Arrow's impossibility theorem
Lützen, Jesper

Published in:
Historia Mathematica

DOI:
10.1016/j.hm.2018.11.001

Publication date:
2019

Document version
Peer reviewed version

Document license:
CC BY-NC-ND

Citation for published version (APA):
How Mathematical Impossibility Changed Welfare Economics.

A History of Arrow’s Impossibility Theorem

Abstract

During the 20th century, impossibility theorems have become an important part of mathematics. Arrow’s impossibility theorem (1950) stands out as one of the first impossibility theorems outside of pure mathematics. It states that it is impossible to design a welfare function (or a voting method) that satisfies some rather innocent looking requirements. Arrow’s theorem became the starting point of social choice theory that has had a great impact on welfare economics. This paper will analyze the history of Arrow’s impossibility theorem in its mathematical and economic contexts. It will be argued that Arrow made a radical change of the mathematical model of welfare economics by connecting it to the theory of voting and that this change was preconditioned by his deep knowledge of the modern axiomatic approach to mathematics and logic.

Introduction

Arrow’s impossibility theorem was probably the first mathematical impossibility theorem that had a great impact on the social sciences. The theorem and the accompanying novel mathematical approach to welfare economics was a result of Arrow’s radical rejection of the cardinal measure of utility and led to a whole new branch of welfare economics called social choice theory. With his new mathematical model Arrow placed social choice theory in a tradition of discussions concerning methods of voting.

Despite Baujard’s claim that “the history of welfare economics is hardly known and studied” (Baujard 2014, 1) some historians of economics including Baujard herself, have investigated this subject from economic and philosophical perspectives. Moreover, the history of the theory of voting has been the subject of several papers and books (see e.g. McLean and Urken 1995). Yet, there is no comprehensive historical account of Arrow’s impossibility theorem considered in the contexts of both welfare economics and the theory of voting, two contexts that it helped bring together. The present paper attempts to give such a connected account. Moreover, other accounts of Arrow’s impossibility theorem focus on its implications for economics and the theory of voting and downplay the mathematical content of the theorem. This paper is conceived as a contribution to the history of mathematics. The contexts of voting and welfare economics are important for the historical analysis, but the main focus is on the development of Arrow’s new mathematical formalism (or model), how it involved what appears to be the first impossibility theorem in the social sciences, and how it led to a new branch of welfare economics: Social choice theory.

The paper is structured as follows: In the first half of the paper we shall follow Arrow’s discovery of his impossibility theorem presenting the earlier ideas and developments on the way as he encountered them. Then comes an account of the novel ideas introduced by Arrow including a sketch of his proof of the impossibility theorem followed by a section on the reception of Arrow’s new approach to the theory of social choice. The paper concludes with some philosophical reflections and a discussion of the driving forces behind the historical development.

Arrow’s encounter with logic and order relations
When Kenneth J. Arrow (1921-2017) was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 1972, he composed a short autobiography (Arrow 1972a) and since then he has given several accounts of the developments that led him to the impossibility theorem (Arrow 1983, 1986, 1991, 2014, 2016 as well as Kelly and Arrow 1987). Although the details vary (but in a consistent way) the accounts all share a rhetoric of destiny: the ideas forced themselves on him despite his own reluctance. According to Arrow his other economic works started from an accepted tradition and were pursued in a deliberate fashion. “My work on social choice, however, did not come out of prolonged scrutiny of a previously recognized problem. It seemed to be more a concept that took possession of me – and had been trying to for some time” (Arrow 1983, 1).

Arrow became interested in subjects related to his impossibility theorem as a child during the Democratic Convention in the U.S. in 1932 when he and his sister followed the fate of the candidates on the radio. However, according to Arrow his interest in elections was not extraordinary and it did not steer him toward the impossibility theorem (Arrow 2014). His undergraduate studies of mathematics, on the other hand, had a direct influence on his later formulation of the welfare problem. In particular he was “fascinated by mathematical logic, a subject I read on my own until, by a curious set of chances the great Polish logician, Alfred Tarski, taught one year at The City College (in New York), where I was a senior. He chose to give a course on the calculus of relations, and I was introduced to such topics as transitivity and orderings” (Arrow 1991, 84).

Axioms for ordered sets had been formulated more than half a century earlier in connection with the construction and axiomatic description of the real numbers (Dedekind 1872 and Hilbert 1900). In set theory and logic, orderings had become important through the investigations of Cantor’s famous well-ordering theorem. In this connection, Ernst Schröder (1890 – 95) had studied orderings and Felix Hausdorff (1914) had formulated the axioms of totally and partially ordered sets. They were studied in more generality in Alfred Tarski’s (1901-1983) Einführung in die mathematische Logik (1937) and in Garrett Birkhoff’s (1911-1996) book Lattice theory from 1940.

Thus, as pointed out by Arrow, the idea of an abstract ordered set was in the 1940s “familiar in mathematics and particularly in symbolic logic” (Arrow 1950, 331). Arrow himself learned the subject so well that he was given the task of proofreading the English translation of Tarski’s Introduction to Logic that came out in 1941 (Tarski 19411, Supes 2017). The axiomatic approach to ordered sets also made its way into an economic text, namely von Neumann and Morgenstern’s book Theory of Games and Economic Behavior (von Neumann and Morgenstern 1944, p. 26, 589). Still, Arrow was probably right when, in 1950 he observed that the “notation [was] not customarily employed in economics” (Arrow 1950, 331)

**Arrow’s encounter with economics. The cardinal and ordinal view of welfare economics**

Arrow was introduced to mathematical economics during his graduate studies at Columbia University. In particular he was greatly influenced by the mathematical statistician and economist Harold Hotelling:

> I 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

1 “I also owe many thanks to Mr. K.J. Arrow for his help in reading proofs” (Tarski 1941, preface xvi).
The distinction between a cardianalist and an ordinalist approach to welfare economics was to become a central element of Arrow’s reformulation of the mathematical model of welfare economics and his impossibility theorem. Welfare economics was (and still is) a mixture of ethics, political theory and economics whose roots can be traced back to 18th century utilitarianism, in particular to the work of Jeremy Bentham (1748-1832) according to whom “it is the greatest happiness of the greatest number that is the measure of right and wrong” (Bentham (1776). Preface, 2nd para.). Although this ethical rule was criticized by Kant it became an important guide for political economics in the 19th and 20th century. It was usually interpreted in the sense that “utility” must be maximized. In the 19th and early 20th century, utility of a particular product or more generally of a particular social state for a particular individual was often believed to be measurable by a real number (cardinally) for example in terms of money. Moreover, it was usually assumed that one could add (or in other ways compound) individual utilities to obtain a total utility. The ethically correct social state would then be the state that maximizes this total utility.

As mathematical analysis was introduced in economics by Antoine Augustin Cournot (1801–1877), Léon Walras (1934-1910) and Francis Ysidro Edgeworth (1845-1926) liberal economists tried to prove that the total utility would be maximized in a free market. This work constitutes what Samuelson has called the “old welfare economics” (Samuelson 1947, 249). The “new welfare economics” began to emerge around 1930 when economists began to question the reasonableness of the concept of a total (or societal) utility: 1. Is it possible to measure the utility of a particular social state for an individual? and 2. Is it possible to compound individual utilities to yield a total societal utility? Bentham himself had previously expressed his doubts concerning the last mentioned problem:

Tis in vain to talk of adding quantities which after the addition will continue distinct as they were before, one man’s happiness will never be another man’s happiness: a gain to one man is no gain to another: you might as well pretend to add 20 apples to 20 pears… (quoted from Arrow 1951, 11).

Later it was also pointed out that one could compound individual utilities in many ways: one could add them, as suggested here by Bentham but one could also multiply them or add some function of the individual utilities. The resulting social optimum may very well turn out to be different under the different ways of compounding.

So Economists tried to deduce what one could conclude if one assumed that individuals were able to order social states but without necessarily being able to measure their utility cardinally and without assuming that one could compound individual utilities to obtain a societal utility (the ordinal approach). One of the steps along the way was taken already by Vilfredo Pareto (1848–1923) in (1910). He defined an economic system to be in an optimal state (Pareto optimum) if one cannot make any one individual better off without making at least one other individual worse off. This concept of an optimum does not require an interpersonal comparison of utility nor a cardinal measure of individual utility. Pareto and his successors tried to prove that under certain assumptions a free market will lead to a Pareto optimum. The rigorous proof was given in 1954 by Arrow and Gérard Debreu. However, the introduction of the concept of a Pareto
optimum does not solve the problem of social choice because there are typically many different Pareto optima many of which will be ethically problematic. For example, a Pareto optimum can be at a point where a few individuals are extremely rich and all other individuals of society are very poor. Moreover one still needed a criterion to decide which of the Pareto optima should be preferred by society.

Another concept that was introduced to avoid the assumption of utility as a cardinal measure was the concept of indifference curves. Instead of measuring the absolute utility of a commodity it compares the utility of two commodities $x$ and $y$. The amount $X$ of $x$ and the amount $Y$ of $y$ is represented in a plane with $X$ and $Y$ along the axis. Then one draws a curve through those points the combination of whose $x$ and $y$ have the same utility for a given person. These curves are the indifference curves. They can replace cardinal utilities in many arguments.

These attempts at avoiding intrapersonal comparisons of utility were synthesized and clarified mathematically by the American economist Abram Bergson (1914-2003) in his paper *A Reformulation of Certain Aspects of Welfare Economics* (1938). His ideas were in turn extended and further clarified in the classical treatise *Foundations of Economic Analysis* (1947) by his friend and colleague Paul A. Samuelson (1915-2009). These classical treatises presented what Samuelson called the new welfare economics.

From the beginning, mathematical economics was a battlefield. Non-mathematically inclined economists were highly critical of the mathematical approach, and mathematical economists criticized each other’s analyses for unclear concepts, faulty or unrigorous arguments or incorrect interpretations of the mathematical results. This critical friction holds true of welfare economics as well. Thus in his historical account of welfare economics Samuelson (1947, 203-228) was highly critical of most of his predecessors in particular their sloppy use of infinitesimals. A typical remark by Samuelson reads: “And if Marshall did arrive at conclusions which are not completely wrong, it is nevertheless clear that he arrived at them for the wrong reasons” (Samuelson 1947, 207). Only Bergson was praised unconditionally by Samuelson.

The basic concept of Bergson’s and Samuelson’s “new welfare economics” was the social welfare function. This function which replaced the old sum of the individual utilities was supposed to measure the total welfare in society. It was a function of many real variables that represent amounts of consumer goods, amounts of labor, and other variables (Bergson 1938, 312, Samuelson 1947, 221). It is less clear what the range of the welfare function is:

The function need only be ordinally defined, and it may or may not be convenient to work with (any) one cardinal index or indicator.... Utilizing one out of an infinity of possible indicators or cardinal indices, we may write this function in the form

$$ W = W(z_1, z_2, \ldots) $$

where the $z$’s represent all possible variables, many of them non-economic in character (Samuelson 1947, 221).

---

2 Bergson distinguished between the welfare function $W$ and the economic welfare function $E$. This distinction is irrelevant in this paper.
Whenever Bergson and Samuelson work mathematically with the welfare function they work with a real valued function, i.e. what Samuelson called a cardinal indicator. But they tried as far as possible to derive consequences that are independent of the cardinal indicator and rely only on the ordering of the values of the welfare function.

The goal of the “new” welfare economics was to maximize the welfare function or one of its cardinal indicators $W$. According to Samuelson this problem is easy in the sense that it can be analyzed by mathematical means: “Once there has been specified a clearly defined $W$ function, the final optimum point can be easily determined” (Samuelson 1947, 236). The methods used to determine the optimum and to deduce economically interesting properties about the optimum are the classical methods of mathematical analysis developed in the 18th and 19th century: Partial differentiation, Lagrange multipliers etc. that implicitly require the functions to be once or twice differentiable.

The subject could end with these banalities were it not for the fact that numerous individuals find it of interest to specialize the form of $W$, the nature of the variables, $z$, and the nature of the constraints (Samuelson 1947, 222).

So what was less trivial was to come up with an acceptable welfare function. This is not primarily a mathematical but an ethical problem. Samuelson formulated 8 different “assumptions” that welfare economists have made about the nature of the welfare function. Some of them are of a rather technical nature, such as the first: “prices are not usually included in the welfare function itself”. Others have stronger philosophical implications, such as in the 5th assumption:

A more extreme assumption, which stems from the individualist philosophy of modern Western Civilization, states that individuals’ preferences are to “count”. If any movement leaves an individual on the same indifference curve, then the social welfare function is unchanged, and similar for an increase or decrease (Samuelson 1947, 223).

The seventh assumption involves an even more “controversial value judgment”:

It is that the welfare function is completely (or very nearly) symmetrical with respect to the consumption of all individuals (Samuelson 1947, 224).

While engaging in a critique of these different assumptions Samuelson derived some of their consequences for welfare economics. He could derive many necessary conditions for the optimum but these would be satisfied by many social states or points. “Without a well-defined $W$ function, i.e., without assumptions concerning intrapersonal comparisons of utility, it is impossible to decide which of these points is best” (Samuelson 1947, 244).

So Bergson’s and, in particular, Samuelson’s “new” welfare economics were based on a welfare function that was in principle defined only ordinally. But the only mathematical way they could represent it was through a cardinal indicator, i.e. through a real valued function that made it possible to use classical analysis.

While Bergson and Samuelson tried to avoid the cardinal interpretation, Von Neumann and Morgenstern argued in their book *Theory of Games and Economic Behavior* (1944) that “treating utilities as numerically
measurable quantities is not quite as radical as is often assumed in the literature” (von Neumann and Morgenstern 1944, 16). In fact they argued for a cardinalist point view in several ways, using analogy with the mathematization of other fields such as the theory of heat and other methodological arguments. Their strongest argument was purely mathematical. They assumed that an individual can order not only all pure alternatives A and B but also combinations of two alternatives $pA + (1-p)B$ meaning situations where there is a probability p of getting A and 1-p of getting B. Allowing comparisons of such combined states they could show that there is a cardinal utility function describing the situation and it is unique up to a linear (affine) transformation. In the first edition of their book they indicated the argument on p. 15-31 and in the second edition they added a more formal proof in an appendix.

In his work on the impossibility theorem Arrow sided with Bergson and Samuelson’s ordinalist approach but radicalized it by rejecting any use of a cardinal indicator and thus any use of mathematical analysis. His recollections quoted at the beginning of this section suggests that in 1941 when he obtained his master’s degree from Columbia University he already considered it “obvious” that one could use abstract order relations to describe the ordinal version of consumer’s and political choice. It may have been “obvious enough” to Arrow, but as we have seen, no other mathematical economist had seen how abstract order relations could help avoiding any use of cardinal indicators. Indeed, Arrow was one of the few mathematical economists who had a training that prepared him to make the connection between economics and the axiomatic approach to ordered sets. And at first he did nothing with the observation.

The Condorcet Paradox

After graduating from Columbia University Arrow joined the war efforts and worked for four years as a weather officer in the United States Army Air Forces. Then he spent a very creative three year period 1946-49 partly as a PhD student at Columbia University and partly as a research associate at the Cowles Commission for Research in Economics at the University of Chicago. In connection with his eventually abandoned work on a dissertation on subjects concerning business economics he was confronted with the so-called Condorcet Paradox in 1946 or 1947 (Arrow 1991, 84; 1983, 2):

I had observed that large corporations were not individuals but were supposed ... to reflect the will of their many stockholders. To be sure, they all had a common aim to maximize profits. But profits depend on the future, and the stockholders might well have different expectations as to future conditions. Suppose the corporation has to choose among alternative directions of investment. Each stockholder orders the different investment policies by the profit he or she expects. But because different stockholders have different expectations, they may well have different orderings of investment policies. My first thought was the obvious one suggested by the formal rules of corporate voting. If there are two investment policies, call them A and B, that one chosen is the one that commands a majority of the shares (Arrow 1986, 48).

When there were more than two possible investment policies

the idea seemed natural to me to choose the one that would get a majority over each of the other two. But I found an unpleasant surprise. It was perfectly possible that A has a majority against B, and B against C, but that C has a majority against A, not A against C..... The observation struck me as one that must have been made by others, and indeed I wondered if I had heard it somewhere. I still don’t
know whether I did or not. In any case the effect was rather to cause me to drop the whole matter and study something else (Arrow 1986, 48-49).

In fact, Arrow’s unpleasant surprise had been experienced many times before. It is now named after Nicolas de Condorcet (1743-1794) who discussed it in his *Essay on the application of analysis to the probability of decisions made on a majority vote* (1785). This famous essay was written as a tribute to Condorcet’s mentor and friend the liberal economist Jacques Turgot (1727-1781). According to Condorcet’s preface Turgot was persuaded that the truths of the moral and political sciences are susceptible to the same certainty as those that form the system of physical sciences and even as the branches of the sciences which, as astronomy, appears to approach mathematical certainty” (Condorcet 1785, i).

With his essay Condorcet wanted to show by an example that Turgot was right. He addressed the following problem: How can a group of people (a jury or a nation) decide what is true when the individual members of the group do not have certain knowledge? More precisely, can one calculate the probability that majority decisions are true, or at least not incorrect? This was an urgent question a decade after the independence of the United States of America and closely connected to Condorcet’s lifelong work in the service of the enlightenment.

In accordance with his enlightenment convictions, Condorcet believed that there was a true answer to all questions, even questions in the social sciences. He attributed to each voter a probability that he would vote for the truth. This probability would depend on the level of enlightenment of the voter, being close to one for highly enlightened persons and below ½ for unenlightened persons. In the first part of his essay Condorcet considered an assembly of voters with a given distribution of such personal probabilities and determined the probability that the majority would vote for the truth. He first showed how Bernoulli’s law of large numbers could solve the problem when there are two alternatives to choose from.

For the purpose of the present paper Condorcet’s subsequent discussion of situations with three or more alternatives is more interesting because here he pointed out a dilemma. If a proposition is a combination of two simple propositions it can happen that the combination that obtains the most votes is not a combination of the two simple propositions that have received most votes. Condorcet gave the following example:

Assume that the proposition is a combination of the two simple propositions $A$ and $a$ with their negations $N$ and $n$. Then there are four possibilities for the combined proposition: $(A,a)$, $(A,n)$, $(N,a)$ and $(N,n)$ .

Condorcet considered an assembly of 33 persons whose votes are distributed as follows:

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>

Here the combined proposition $(A,a)$ gets the highest number of votes, but as Condorcet pointed out the alternative $a$ only got 14 votes compared to 19 for $n$. “Thus it is the two propositions $A$ & $n$ that should be chosen and the second opinion rather than the first that in reality has the majority.” (Condorcet 1785, xlviij)
From this observation Condorcet drew the general rule that in order to decide the result of a composed proposition one has to decompose it into simple choices between propositions and their negations (Condorcet 1785, xlviij). The combined outcome of the vote should then consist of the combination of the simple propositions that obtains the majority. If the combined proposition consists of \( n \) simple propositions, there can be up to \( 2^n \) different possibilities for the combined choice, but in many cases some of them are inconsistent and must therefore be left out.

However, Condorcet discovered that even if the voters do not vote for inconsistent combinations of the simple propositions, it can happen that the voting leads to an inconsistent result. His famous examples concern the election between three candidates \( A, B \) and \( C \) (Condorcet 1785, lvj-lxx). He considered such an election as a combination of the following three simple alternatives: \(^3\) \( A>B \) (and its negation \( B>A \)), \( A>C \) (and its negation \( C>A \)) and \( B>C \) (and its negation \( C>B \)). A vote consisting of a ranking of the three candidates will be a combination of the three simple alternatives. There are \( 2^3 \) combinations, but the following two combinations are contradictory (intransitive in modern parlance): (\( A>B, B>C, C>A \)) and (\( B>A, A>C, C>B \)).

Assume first with Condorcet that there are 60 voters the remaining possible combinations get the following number of votes:

- \( A>C>B \): 23 votes
- \( A>B>C \): 0 votes
- \( B>C>A \): 19 votes
- \( B>A>C \): 0 votes
- \( C>B>A \): 16 votes
- \( C>A>B \): 2 votes.

The most popular combined vote is the first one that has \( A \) as the winner. However, according to Condorcet we must look at the three simple votes that make up the combined votes. And here out of the 60 votes the following holds:

- \( B>A \) has the majority of 35 votes over 25 for \( A>B \)
- \( C>A \) has the majority of 37 votes over 23 for \( A>C \)
- \( C>B \) has the majority of 41 votes over 19 for \( B>C \)

Thus \( C \) has a majority over both \( A \) and \( B \) and is therefore the winner according to Condorcet, although in the combined votes, only 18 had \( C \) in the first place compared to 19 with \( B \) in the first place and 23 with \( A \) in the first place.

Thus the candidate who was in reality wanted by the majority was precisely the one who according to the ordinary method had the smallest number of votes (Condorcet 1785, lx).

---

\(^3\) I use the notation \( A>B \) for Condorcet’s “\( A \) is valued higher than \( B \)”. 
So what Condorcet suggested was to think of the election as divided up into pairwise majority votes. The candidate who has a majority over each of the other candidates is to be considered the winner. Today we call such a winner a Condorcet-winner, and it is often argued that such a winner has a good claim to be considered the true winner.

In the example above, Condorcet’s method of pairwise majority gave a clear winner, although another winner than the usual plurality method. However, Condorcet also gave the following famous variation of the above example where his method did not immediately yield a winner.

A>C>B: 23 votes
A>B>C: 0 votes
B>C>A: 17 votes
B>A>C: 2 votes
C>B>A: 8 votes
C>A>B: 10 votes.

In this case the distribution of the votes in each of the simple alternatives is as follows:

A>B has the majority of 33 votes over 27 for B>A
C>A has the majority of 35 votes over 25 for A>C
B>C has the majority of 42 votes over 18 for B>C

Thus, whomever of the three candidates you chose there is another candidate who is preferred by the majority of the voters. The result is the intransitive combination: A>B>C>A called “inconsistent” by Condorcet. This is the first discovery of the disquieting phenomenon of cyclical majorities (to use a phrase introduced by Duncan Black): Although each voter has ordered the candidates in a transitive way, the method of pairwise majority can lead to a non-transitive ordering of the candidates, containing one or more cycles. This phenomenon is often called the Condorcet paradox.

A profusion of voting methods

Already in 1770, before the appearance of Condorcet’s Essay, his compatriot the military mathematician Jean-Charles de Borda (1733 – 1799) had discovered the inconvenience of the ordinary plurality voting system where each voter votes for one candidate and the candidate with most votes win. Instead, he suggested the so-called Borda count (or rank order) in which the candidates are not only ordered by the voters but are ranked by each voter who assigns a decreasing number of points to the candidates. If there are n candidates the first on a ballot is assigned n points by the voter, the second n-1 points and the last 1 point. The number of points allotted by the voters to each candidate are added and the winner is the candidate who gets the highest number of points. Condorcet admitted that this method has the advantage over his own method that it will always pick a winner (except when two candidates get the same number of points). However, he also showed by an example that the method may elect a person B, who with his own
method of pairwise majority would have lost to another candidate A, and even such that “by preferring B to A one prefers the one whose probability of merit is not only below the probability of the other, but a probability below ½ to a probability which is above ½”. [Condorcet 1785, clxxvij]. Borda admitted that the use of the arithmetic progression $n, n-1, n-2,...,3,2,1$ in assigning points to the candidates is somewhat arbitrary. Any decreasing progression would reflect the order of priorities of the voters. However, he argued that since a ranked list does not carry any information about how much each candidate is preferred over the candidate above and below him, it is most rational to assume that the distance between each candidate is the same. In 1795 Laplace gave a more sophisticated probability argument in favor of the Borda count (see Black 1998, 213-215).

Recently it has been discovered that the Borda-count and the Condorcet majority method had already been put forward by medieval thinkers. In fact, in connection with questions of elections of abbots or abbesses, popes and emperors, Ramon Llull suggested the majority rule of pairwise voting (1283) and in 1433 Nicolas of Cusa recommended the Borda count (McLean 1990, Hägele and Pukelsheim 2001, Mclean, Lorrey and Colomer 2008). But apparently the medieval authors did not discover the Condorcet paradox.

During the 19th and 20th century many voting methods were developed and many of these methods were used in practice in elections on all levels from national elections to votes in small clubs and societies. There is a rich literature on electoral (or voting) systems and its history (see O'Connor and Robertson 2017). Here it will suffice to notice that the development was not continuous and accumulative. Many ideas, methods and critical reflections were discovered and rediscovered over and over again. The reason was partly that the problem of voting did not belong to one clearly defined subject area. People of many different trades contributed: Statesmen such as Alexander Hamilton and Thomas Jefferson, architects such as William Robert Ware, lawyers such as Andrew Inglis Clark and Thomas Hare and mathematicians such as Edward J. Nanson, Carl Andræ, Thorvald Nicolai Thiele and Charles Dodgson. Many contributions to the subject of voting systems offered a method and argued for it on various grounds, while showing the undesirable consequences of other methods. Some authors like Dodgson and Thiele, provided more systematic analyses of different properties of voting systems.

Charles Dodgson (1832-1898), better known under his pen name Lewis Carroll, was probably the most studied 19th century contributor to the theory of voting. And yet, his writings on the subject were limited to three pamphlets (1873-76), one cyclostyled sheet (1877), nine letters in a newspaper (1881-1885) and one booklet: *The Principles of parliamentary Representation* (1884), the latter only dealing with elections with two political parties. A planned longer book on voting never saw the light of day (Black, 1998, 241). Many of Dodgson’s contributions were written in relation to elections at Christ Church College, Oxford where he was a mathematical lecturer. They did not reach a wide circulation but were rediscovered by Duncan Black (1908-1991) who included an analysis of them and reprinted the pamphlets in the second part of his main work: *The Theory of Committees and Elections* (1958). In his pamphlets Dodgson suggested several methods of voting, one similar to the Borda count, one using the Condorcet method (when possible), several employing methods of elimination and one using a method of marks, where each voter is given a certain fixed number of marks that he can distribute at his own will to the various candidates. Several of his procedures were combinations of these and other methods.

---

4 For a history with sources of voting theories before 1900 see (McLean and Urken 1995).
At least two of the contributors to the 19th century theory of voting were aware of the Condorcet paradox: Dodgson and Nanson (1850-1936). Black convincingly argued that Dodgson discovered the paradox independently of Condorcet and never learned of his 18th century predecessor (Black 1998, 227-228). Nanson (1882), on the other hand, referred to Condorcet.

Like Black’s other works, his historical research on Dodgson had limited impact at first. When Black died he left a lot of manuscripts for a never published book on Dodgson’s approach to proportional representation. These manuscripts were edited by McLean, McMillan and Monroe and published as a book (Black 1996). As a result many historians now agree with Black that Dodgson’s understanding of the subject (in particular the effects of circularity or intransitivity) “was second only to that of Condorcet”.

As we saw above, Arrow discovered the Condorcet paradox independently in 1946 or 1947 although he had a feeling that he may have heard about it earlier. Subsequently he became aware of Condorcet’s contribution and in the second edition (1963) of his book Social Choice and Individual Values he explicitly referred to Condorcet’s Essay.


When Arrow rediscovered the Condorcet paradox in 1946-47 his first reaction was to “drop the whole matter and study something else” (Arrow 1986, 49). However, he was subconsciously drawn back to the problem of voting one year later when he was in Chicago at the Cowles Commission.

About a year later my thoughts recurred to the question of voting, without any intention on my part. I realized that under certain special but not totally unnatural conditions on the voters’ preferences, the paradox I had found earlier could not occur. This I thought worth writing about. But when I started to do so I picked up a journal and found the same statement in an article by an English economist Duncan Black. The result that Black and I had found could have been thought of any time in the last one hundred and fifty years. That two of us came to it at virtually the same time is an occurrence for which I have no explanation. Priority in discovery is a spur to science, and being anticipated was correspondingly frustrating. I again dropped the study of Voting (Arrow 1986, 49).

The condition (re)discovered by Arrow was the so-called condition of single peaked preference curves originally found in 1942 by Duncan Black who was in fact not English as stated by Arrow but born in 1908 in Scotland. He was the first writer who devoted the majority of his career studying voting methods. After having completed a master’s degree in mathematics at Glasgow University he enrolled for a second degree in economics and politics. He graduated for the second time in 1932 and began to teach economics first at the Dundee School of Economics and later at University College of North Wales (now Bangor University) and Glasgow University (McLean, McMillan and Monroe 1998). While in Dundee the work on the nature of firms or companies by his colleague Ronald. H. Coase and the Italian school of writers on public finance induced him to try to formalize a pure science of politics, which he considered to be equal to a science of economics:

---

5 According to (Arrow 1983, 3) it happened in a context of political parties “arrayed in a natural left-right ordering.”
6 According to (Kelly and Arrow 1987, 50) the paper he picked up was (Black 1948).
The main reason which I can give in substantiation of this view (that a science of politics is of the same kind as a science of economics) is that it is possible, using terms which are precisely those of economic Science, to construct a Theory of Committee Decisions. In getting a theory of the committee, however, it is clear that we at the same time get a sufficient means to construct a Theory of Politics (Black 1998, 353).

The theory that he began developing in 1934 dealt with the shape of decisions taken by concrete government bodies. The great breakthrough came “in a flash” (Black 1998, 388) in February of 1942 during a night while he was fire-watching in the green drawing-room of Warwick Castle:

Acting apparently at random, I wrote down a single diagram and saw in a shock of recognition the property of the median optimum. This could be got by interpreting the diagram I had drawn, in terms of a committee using a simple majority, whose members’ preferences in regard to the motions put forward, could be represented by a set of single-peaked curves (Black 1998, 389).

In his book *The Theory of Committees and Elections* (Black 1958) he explained the result as follows: He illustrated the individual committee member’s orderings of a finite number of motions in a diagram with the motions (or their indices) along the first axis and the ordinates representing the preferences, such that a motion with a larger ordinate is preferred to one with a lower ordinate. Connecting the ordinate points by straight line segments he obtained a preference curve (see Figure 1).

![Figure 1: The top figure is a modern graph representing three voters and five options. The rankings are single-peaked. Option D is the median and thus the Condorcet winner. The bottom figure is Fig. 18 in Black’s book of 1958. One must assume that the curves follow the first axis where it is not drawn so that the individual put all those alternatives on the same level at the bottom of their preference scale.](image-url)
If a preference curve has a single peak from which it slopes downwards in both directions Black called it single-peaked: “a single-peaked curve is one which changes its direction at most once, from up to down” (Black 1958, 10). Whether a preference curve comes out single-peaked or not depends on the ordering of the motions along the first axis. If it is possible to find an ordering of the motions such that all committee members’ preference curves are single-peaked Black says that all the curves are single-peaked. Denoting the peaks (maxima) of the preference curves $O_1, O_2, ..., O_{2n+1}$ ordered in increasing order along the first axis (assuming that there are an odd number of members), Black easily deduced that the median motion $O_{(n+1)}$ “can get at least a simple majority against every other, and it is the only value which can do so” (Black 1958, 21). In other words, the median is a Condorcet winner in this case. This is one of his most famous results, and the one that he realized while fire-watching. The celebrity of the result is not due to the difficulty of its proof. In fact Black’s proof was easy. Indeed, the median option (in cases where there are an odd number of options) will beat all the options to the left of it, because the voter who has put it in the first place, as well as the voters having their first choice to the left of him, will give the median option more votes than the options to the right of it. And they constitute a majority. Similarly for the options to the left of the median option. Black argued that single-peakedness was not as uncommon as it might seem:

It (single-peakedness) would be particularly likely to happen if the committee were considering different possible sizes of a numerical quantity and choosing one size in preference to the others. It might for example be reaching a decision with regard to the price of a product to be marketed by a firm, or the output for a future period, or the wage rate of labour, or the height of a particular tax, or the legal school-leaving age, and so on. In such cases the committee member in arriving at an opinion on the matter, will often try initially to judge which size is for him the optimum. Once he has arrived at his view of the optimum size, the further any proposal departs from it on the one side or on the other, the less he will favour it. The valuations carried out by the member would then take the form of points on a single-peaked or \(\cap\)-shaped curve (Black 1958, 12).

When Black had discovered that “the political problem had been transformed into a mathematical problem” (Black 1958, 389), he soon developed the theory further:

In the early months, working on the geometrical version of the theory, I took for granted that with a simple majority in use, the answer, irrespectively of the shapes of the preference curves, would be determinate (Black 1998, 390).

However, he was soon surprised to discover the problem of intransitivity that had been discovered earlier by Condorcet. He later described his discovery in vivid emotional terms:

Later in working out an arithmetical example, an intransitivity arose, and it seemed to me that this must be due to a mistake in the arithmetic. On finding that the arithmetic was correct and the intransitivity persisted, my stomach revolted in something akin to physical sickness. Not only was the problem to which I had addressed myself more complicated than I had supposed, it was of a different kind. (Black 1991, 390)

Black went on to investigate the problem of intransitivity in more detail, asking and partly resolving questions like: How probable is it that intransitivity occurs? What are the properties of the cycles that lead to intransitivity. What happens if the vote mixes two or more single-peaked questions, etc. He composed several papers and a book on this “Pure Science of Politics,” but according to his friend and colleague Ronald H. Coase (1910-2013) (1998) the book was rejected by four British publishers and the papers were rejected by the leading British economic journals. Still, in 1948-49 he succeeded in getting 6 papers
published in foreign journals.\textsuperscript{7} He also composed a paper with his colleague R.A. Newing and sent it to the leading American journal Econometrica in November, 1949 (Coase 1998, xi). After one and a half years of delay the editor of the journal answered:

\begin{quote}
I would like to state that if the interrelationship with Arrow’s recent monograph could be brought out clearly throughout the paper, I would like very much to recommend your manuscript for publication in Econometrica (quoted by Coase 1998, xii).
\end{quote}

This answer “shocked and angered Black” (Coase 1998, xiii). Not only was Black asked to relate to Arrow’s ideas that were developed later than his own, but he also realized that the managing editor of Econometrica was serving as assistant Director of Research at the Cowles Commission where Arrow had written his book. It seemed as if Black’s paper had been deliberately postponed in order to give Arrow priority.\textsuperscript{8}

Black and Newing withdrew their paper from Econometrica and succeeded in having it published as a separate booklet in 1951.

The effect on Black of this decision was in fact profound. He became suspicious of the motives of others and retired within himself, largely devoting himself henceforth in his researches to an historical study of the thought of his predecessors, above all, Lewis Carroll (Coase 1998, xiv). In 1958 Cambridge University Press published his book on pure science of politics under the title \textit{The Theory of Committees and Elections} including his historical research as part 2. If the book had appeared in the late 1940s it would surely have had a greater impact than it had in the late 50s after the publication of Arrow’s research. Black published very little after 1958 but still towards the end of the 20\textsuperscript{th} century he was “recognized as a founding father of theoretical politics”.\textsuperscript{9} Thus a volume of papers “in honor of Duncan Black” was published in 1981 by G. Tullock. It included a biographical sketch of Black’s life by R.H. Coase. Republications of his most important works including an unpublished book on Dodgeson’s theories of proportional representation with fine and laudatory comments by the editors McLean, McMillan and Monroe has recently added to Blacks fame (cf. Black 1996 and 1998). We shall return to Black’s evaluation of Arrow’s results after we have discussed them in the following sections.

\textbf{Arrow’s discovery of his Impossibility theorem}

Until 1948 Arrow had unknowingly walked in the footsteps of Black, only difference being that Arrow discovered the Condorcet paradox first and single-peakedness afterwards, where Black discovered the two phenomena in reverse order. Where Black had been frustrated by his inability to get his ideas published in the best journals, Arrow was frustrated to discover that Black had anticipated him.

Again, Arrow did not pursue the matter, not only because he was frustrated by Black’s priority but also “because I felt all this was distracting me from what I took to be my obligation to serious economic work, specifically using general equilibrium theory to develop a workable model as a basis for economic analysis” (Arrow 1991, 85).

But it did not last long until Arrow returned to the subject and this time he went in a different direction from Black. The crucial turn of events happened in the summer of 1948 when Arrow

\textsuperscript{7} See references in Black 1998 (410-411) and in (Arrow 1951, p. 6).
\textsuperscript{8} Arrow did not mention Black in his first paper on the impossibility theorem (Arrow 1950) but in the subsequent monograph (Arrow 1951, 6 and 75-80) he referred to Black’s publications and analyzed his result on single peaked preference curves redefining the condition of singlepeakedness in his own formal way.
\textsuperscript{9} Quoted by the editors of (Back 1996, xiv) from an anonymous obituary of Duncan Black in Journal of Theoretical Politics 3(3) (1991), 276.
was invited to the then-new RAND Corporation, which was trying to develop game theory as a tool for analysis of international relations and military conflict. During one of the Coffee breaks ... Olaf Helmer\textsuperscript{10} ... told me he was troubled by the foundations of this application. Game theory was based on utility functions for individuals; but when applied to international relations the “players” were countries, not individuals. In what sense could collectivities be said to have utility functions? I told him that ... the question... had been answered by Abram Bergson’s notion of the welfare function (Arrow 1983, 3-4) (quoted in Black 1998, xxxi).

When Arrow, on Helmer’s request, tried to write down the details of Bergson’s theory he was led to his own new conception of social welfare and the social welfare function:

I quickly perceived that the ordinalist viewpoint, which I had fully adopted, implied that the only preference information that could be transmitted across individuals was an ordering. Social welfare could only be an aggregate of orderings (Arrow 1983, 3).

However, when he began to develop this idea he realized that his earlier results had taught him that one could not always derive a preference ordering for a nation from the preference orderings of its citizens by using majority voting to compare one alternative with another.

This left open the possibility that there were other ways of aggregating preference orderings to form a social ordering, that is a way of choosing among alternatives that has the property of transitivity. A few weeks of intensive thought made the answer clear (Arrow 1986, 50).

According to his later recollections (Arrow 1991, 85) he made his discovery in three steps: 1. A few days of trying other aggregation methods made him “suspect that there was an impossibility result”. 2. He found such an impossibility result “very shortly”. 3: “A few weeks later\textsuperscript{11} I made a further strengthening and this was the form expressed above [i.e. Arrow’s impossibility theorem]” (Arrow 1991, 85).

This breakthrough was only possible because Arrow could rely on his thorough knowledge of mathematical logic, and ordering relations in particular: “My studies in logic helped to formulate the question in a clear way” (Arrow, 1986, 50).

The new impossibility result and his new formulation of the problem of social choice became the subject of Arrow’s PhD dissertation, which he defended in 1949. That same year he presented his results at a meeting of the Econometric Society, and the following year they were published as a paper in the \textit{Journal of Political Economy} (Arrow 1950). The year after came his monograph \textit{Social Choice and Individual Values} explaining “the mathematics more completely and adding a number of interpretive comments” (Arrow 1991, 85).

In the 1950 paper and the 1951 book Arrow named his theorem a “general possibility theorem”. In a lecture in the Kenneth J. Arrow Lecture Series, Sen (2014) ascribed this “cheerful name” to Arrow’s “sunny

\textsuperscript{10}The philosopher Olaf Helmer (1910-2011) who worked at the RAND Corporation had studied mathematics and logic in Berlin. As a German Jew he fled to England and in 1937 to the US. Arrow had met him earlier through Tarski whose book on logic Helmer had translated into English (Arrow 2014), (Tarski 1941, xv).

\textsuperscript{11}In (Arrow 1983, 4) Arrow wrote that it took “about three weeks".
"temperament". However, after the lecture Arrow pointed out that it was in fact not he who had decided to use that name. He had always considered himself “a gloomy realist” so instead, someone else, Tjalling Koopmans\textsuperscript{12} insisted on using the word possibility. He was upset by the term impossibility. Now, I cannot say that Tjalling had an extraordinarily sunny disposition either..... But he did dislike the feeling that things could not happen or change. And given that the dissertation was originally posed as a Cowles Commission monograph, I felt that to please Tjalling, I would call it a “possibility theorem”. It was not, however, my idea at all (Arrow 2014, 36).

In the second edition of his monograph (1963) Arrow stuck to his original name but soon it began to be known as an impossibility result. For example Blau (1957) used the name possibility theorem for Arrow’s original formulation of the theorem, that Blau proved to be false, but he called his own slightly rephrased (and correct) theorem an impossibility theorem (Blau 1957, 309). With Kelly’s book Arrow Impossibility Theorems (1978) the name impossibility theorem was canonized and Arrow himself also used this name in later writings.

**Arrow’s new model for social choice. Weak orderings**

As Bergson’s and Samuelson’s welfare economics, Arrow’s theory of social choice dealt with social states that he defined as follows:

The most precise definition of a social state would be a complete description of the amount of each type of commodity in the hands of each individual, the amount of labor to be supplied by each individual, the amount of each productive resource invested in each type of productive activity, and the amounts of various types of collective activity such as municipal services, diplomacy and its continuation by other means, and the erection of statues of famous men (Arrow 1950, 333).

Thus, except for the last witty addition, Arrow considered pretty much the same variables as Samuelson. And as in the earlier welfare economics, Arrow’s social choice theory dealt with ordering of social states, both on the individual level and on the societal level. However, Arrow’s orderings were not represented by a cardinal indicator but was simply defined as a weak ordering or a preference pattern.

It is further assumed that utility is not measurable in any sense relevant to welfare economics, so that the tastes of an individual are completely described by a suitable preference pattern” (Arrow 1950, 331).

Arrow emphasized that individual’s ordering of the social states need not be determined solely by the “commodity bundles which accrue to his lot under each”. Political, ethic and esthetic values may play a role as well: “It is simply assumed that the individual orders all social states by whatever standards he deems relevant” (Arrow 1951, 17).

The mathematical notation and concept used by Arrow to describe preference patterns was the concept that he called a weak ordering of the alternatives. If \(x\) and \(y\) are two alternatives in a set \(S\) of possible

---

\textsuperscript{12} Tjalling Charles Koopmans (1910-1985): Dutch American mathematician and economist. In 1948 he became the director of the Cowles Commission.
alternatives, $xRy$ symbolizes the statement “$x$ is preferred or indifferent to $y$”. He assumed that this relation satisfies the following two axioms:

Axiom I: For all $x$ and $y$, either $xRy$ or $yRx$. (Arrow 1951, 13)

Axiom II: For all $x, y$, and $z$ $xRy$ and $yRz$ imply $xRz$ (Arrow 1951, 13)

He explicitly pointed out that the first axiom implies that $xRx$ for any alternative $x$ and that he did not exclude the possibility that both $xRy$ and $yRx$ could hold. In that case he wrote $xIy$ meaning $x$ is indifferent to $y$. Moreover, “$xPy$ is defined to mean not $yRx$. The statement “$xPy$” is read “$x$ is preferred to $y”’’ (Arrow 1951, 14).

According to Arrow, rationally behaving individuals will make choices that satisfy the two axioms.

Hence one of the consequences of the assumption of rational behavior is that the choice from any collection of alternatives can be determined by knowledge of the choices which would be made between pairs of alternatives (Arrow 1950, 332-333).

Arrow’s use of the new axiomatic structure of an ordered set is a fine illustration of the strength of the modern axiomatic method. In a modern axiomatic structure, the elements are left undefined and therefore the theory can be applied to a great variety of concrete objects, in this case social states. With this in mind, it becomes less surprising that earlier welfare economists had continued to use the real valued measure of utility, although they accepted only an ordinal definition of that concept. The real numbers were for them the most natural ordered set that came to mind. And since they used this richer structure instead of a general ordered set, they also used its other features in particular the ability to use differential calculus. Arrow, on the other hand, denied himself this possibility and therefore had to completely reshape the idea of welfare economics or the theory of social choice.

Arrow also made another independent change in the mathematical model of welfare economics: Where previous authors, including Samuelson, had assumed that the variables entering the welfare function were continuous real variables, and that the welfare function was continuous and even differentiable, Arrow assumed that there was only a discrete set, and actually a finite set of possible social alternatives. This changed the mathematical methods of welfare economics fundamentally from continuous mathematics (analysis) to discrete mathematics.

**Arrow’s Social Welfare Function. The Connection to Voting.**

So according to Arrow each individual $i$ ($i=1,2,...,n$) of the society under investigation chooses an ordering relation $R_i$ that describes his preferences among the possible social states. The problem of welfare economics according to Arrow is to find suitable mechanisms that from the collection of individual

---

13 As we saw above Black also used the ordered straight line as the carrier of the order.
Preference orderings will form a social ordering \( R \) for the society at large. The function that maps the collection of individual orderings into the societal ordering is what Arrow called a social welfare function:\(^{14}\)

**Definition 4:** By a *social welfare function* will be meant a process or rule which, for each set of individual orderings \( R_1, \ldots, R_n \) for alternative social states (one ordering for each individual), states a corresponding social ordering of alternative social states \( R \) (Arrow 1950, 335).

Gone\(^{15}\) is the real valued welfare function whose maximum represents the optimal social state. Instead, the Arrowian welfare function maps a collection (actually an ordered set) of individual orderings of the alternative social states into one societal ordering of the social states. The alternatives ordered at the top of the social ordering are obviously to be considered the optimal social choice (or choices, since there can be several indifferent alternatives at the top). Contrary to later writers Arrow did not designate the social welfare function by any particular letter, but he used the letter \( C \) to denote the social choice function corresponding to a given welfare function. This is a function that maps a subset \( S \) of alternatives (or an “environment” \( S \) in Arrow’s terminology) into the element of \( S \) that the welfare function places highest among the elements of \( S \). So if \( S \) is a subset of the possible social alternatives and \( R \) is a social ordering determined by the social welfare function, Arrow defined:

**Definition 3:** \( C(S) \) is the set of all alternatives \( x \) in \( S \) such that, for every \( y \) in \( S \), \( xRy \) (Arrow 1951, 15).

With this new formulation of the welfare function, the relation to voting became obvious. Where individuals in an election rank candidates, they rank social states in Arrow’s welfare setup. And where it is an election procedure that determines the winner of an election it is Arrow’s new welfare function that does the job in the welfare situation. So, in fact, the theory of voting methods and the theory of social choice were completely merged by Arrow in his new formulation of the latter.

As in Samuelson’s setup, the main question in Arrow’s new formulation was to determine a morally acceptable or a just welfare function. As the earlier works on voting had revealed, this was not a simple matter. As we have seen, many voting methods had been put forward, but all of them had had undesirable properties. Probably inspired by Samuelson, Arrow approached the problem from a new direction. Instead of investigating which properties particular voting methods (or social welfare functions) have, he began with the properties. He proposed a list of conditions loosely corresponding to Samuelson’s assumptions about the welfare function, and then investigated what can be said about social welfare functions that satisfy these conditions. His conditions were formulated in a completely rigorous way in terms of his concept and notation of ordering, and also elucidated in more everyday language.

His conditions are analogous to axioms in a modern mathematical theory. Thus, also in this sense Arrow build on the modern mathematical axiomatic method. Indeed, in this typical 20th century approach to mathematics it is standard to ask for the properties of objects or systems that satisfy certain axioms.

---

\(^{14}\) Arrow modestly wrote: “I will largely restate Professor Bergson’s formulation of the problem of making welfare judgments in the terminology here adopted” (Arrow 1951, 22). We will see that this is a polite understatement of his reshaping of the mathematical model.

\(^{15}\) At least in Arrow’s approach. We shall later see that although Arrow believed he had shown that Samuelson’s “new” welfare economics was not tenable, it continued to be used and developed.
Arrow’s approach to the welfare question and the question of voting is a good illustration of the strength of the new axiomatic method. And it led to the surprising result that no object can satisfy the conditions.

**Arrow’s conditions**

The conditions that Arrow formulated should make sure that the societal ordering would reflect the individual orderings in a way that could be considered just in some way. He did not count the axioms of ordered sets as separate conditions, but he assumed explicitly that both the individual orderings and the societal ordering satisfy the axioms of a weak order relation. In particular the societal ordering must be transitive.

The condition that Arrow numbered as the first concerned the domain of the welfare function. He did not assume that the welfare function was defined for all possible combinations of individual orderings but only for a subset of “admissible” individual orderings. Indeed, there may be a priori reasons for believing that certain orderings are inadmissible.

We will, however, suppose that our a priori knowledge about the occurrence of individual orderings is incomplete, to the extent that there are at least three among all the alternatives under considerations for which the ordering by any given individual is completely unknown in advance (Arrow 1951, 24)

He formulated this condition more formally as follows:

**Condition 1:** Among all the alternatives there is a set $S$ of three alternatives such that, for any set of individual orderings $T_1, \ldots, T_n$ of the alternatives in $S$, there is an admissible set of individual orderings $R_1, \ldots, R_o$ of all the alternatives, such that for each individual $I$, $x \preceq y$ if and only if $x \preceq y$ for $x$ and $y$ in $S$ (Arrow 1951, 24).

This condition will exclude Condorcet’s pairwise majority voting method due to the Condorcet paradox.\(^{16}\)

Since we are trying to describe social welfare and not some sort of illfare, we must assume that the social welfare function is such that the social ordering responds positively to alterations in individual values, or at least not negatively. Hence, if one alternative social state rises or remains still in the orderings 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, 25).

He called this a “positive association of social and individual values” (Arrow 1951, 25).

The condition of “independence of irrelevant alternatives” was formulated as follows by Arrow:

\(^{16}\) In the notes added to the second edition of Arrow’s book from 1963 Arrow actually altered condition 1 to 1’:’” All logically possible orderings of the alternative social states are admissible” (Arrow, 1963, 96). This has since been called the condition of “unrestricted domain”.
Condition 3: Let \( R_1, \ldots, R_n \) and \( R'_1, \ldots, R'_n \) be two sets of individual orderings and let \( C(S) \) and \( C'(S) \) be the corresponding social choice functions. If, for all individuals \( i \) and for all \( x \) and \( y \) in a given environment \( S \), \( x \mathcal{R}_i y \) if and only if \( x \mathcal{R}'_i y \), then \( C(S) \) and \( C'(S) \) are the same (Arrow 1951, 327).

In other words: The social choice made from a subset \( S \) of the possible alternatives is independent of the ranking of the alternatives not in \( S \) or as Arrow put it:

Just as for a single individual, the choice made by society from any given set of alternatives should be independent of the very existence of alternatives outside the given set (Arrow 1950, 337).

He argued for the reasonableness of this condition by pointing out that if the choice function does not satisfy the condition, the death of a candidate in an election would upset the result, even if the deceased candidate had not won the election. He also gave an example to show that the usual Borda count does not satisfy this condition.

In order to make sure “that the individuals in our society are free to choose, by varying their values, among the alternatives available” (Arrow 1951, 28) Arrow formulated the notion of an imposed welfare function:

Definition 5: A social welfare function will be said to be imposed if for some pair of distinct alternatives \( x \) and \( y \), \( x \mathcal{R}_i y \) for any set of individual orderings \( R_1, \ldots, R_n \), where \( R \) is the social ordering corresponding to \( R_1, \ldots, R_n \) (Arrow 1951, 28).

In terms of this new notion, Arrow formulated his 4th condition thus:

Condition 4: The social welfare function is not to be imposed (Arrow 1951, 29).

In other words, if \( x \) and \( y \) are two alternative social states, the voters should be able to vote in such a way that society prefers \( x \) to \( y \). Combined with condition 2 and 3 this means that if all citizens prefer \( x \) to \( y \) then society will prefer \( x \) to \( y \). Arrow called condition 4 “the condition of citizens’ sovereignty”. In contemporary literature it is called weak Pareto efficiency, because it ensures that the social choice is a Pareto optimum.

Arrow’s last condition stated:

Condition 5: The social welfare function is not to be dictatorial (nondictatorship) (Arrow 1950, 339; 1951, 30).

A dictator according to Arrow is an individual whose ordering is picked out by the social welfare function no matter how the rest of the individuals rank the alternatives:

Definition 6: A social welfare function is said to be dictatorial if there exists an individual \( i \) such that, for all \( x \) and \( y \), \( x \mathcal{P}_i y \) implies \( x \mathcal{P} y \) regardless of the orderings \( R_1, \ldots, R_n \) of all individuals other than \( i \), where \( \mathcal{P} \) is the social preference relation corresponding to \( R_1, \ldots, R_n \) (Arrow 1951, 30).

If individual \( i \) is a dictator, his choices will determine society’s choices. The only decisions left to the other individuals are the ranking of the alternatives that the dictator is indifferent about.

As Arrow pointed out all these conditions seem innocent and almost self-evident. They do not assume intrapersonal comparison of the individual’s rankings. For example Arrow did not make any symmetry
assumption similar to Samuelson’s assumption 7. It is not excluded that some individuals’ rankings count more heavily in the formation of the societal ranking than others’ as long as there is no dictator. And yet, Arrow could prove that there does not exist a social welfare function satisfying the conditions.

**Arrow’s Proof of the Impossibility Theorem**

In his book (1951) Arrow first gave a simple proof of the impossibility theorem in the case where there are 2 voters and 3 alternatives. This proof was repeated almost verbatim from the paper of 1950. Then he proceeded to the general case. The general proof relies on the concept of a decisive set that has figured prominently in many later impossibility proofs as well:

Definition 10: The set \( V \) is said to be decisive for \( x \) against \( y \) if \( x \neq y \) and \( xPy \) for all sets of admissible individual ordering relations such that \( xPy \) for all \( i \) in \( V \) (Arrow 1951, 52).

After establishing a strengthening of condition 2, he could prove (consequence 2) that if there exists a set of individual ordering relations \( R_1, R_2, \ldots, R_n \) such that \( xPy \) for all \( i \) in a set \( V \) and such that \( yPx \) for all other \( i \) and such that the social preference relation yields \( xPy \), then \( V \) is decisive for \( x \) against \( y \).

From the condition 4 of citizens’ sovereignty, Arrow then deduced that the entire set of voters is decisive for \( x \) against \( y \) for all \( x \) and \( y \), and further that if a set consisting of a single individual is decisive for \( x \) against \( y \) or \( y \) against \( z \) for specific distinct alternatives \( x, y \) and \( z \), then it is decisive for \( x \) against \( z \). Here he made use of the transitivity of the social ordering. This in turn implies (consequence 5) that a one member set of individuals cannot be decisive for any \( x \) against any \( y \), or said differently, no individual can be a dictator for even one pair of alternatives (the condition of non-dictatorship only stated that no dictator could determine the entire social preference relation).

Armed with these consequences of the conditions satisfied by the social choice function, Arrow could prove that the conditions lead to a contradiction. He assumed that the alternatives \( x, y, z \) could be ordered in any way by the individual voters (condition 1). Since the set of all individuals is decisive for any alternative against any other alternative we can select a set \( V_1 \) with the smallest possible number of members which is decisive for one of the three alternatives \( x, y, z \) over any other of the three (say \( x \) over \( y \)). Let the number of members in \( V_1 \) be \( k \) and denote its members \( 1, 2, \ldots, k \) while \( k+1, \ldots, n \) are the members in its complement \( V_3 \) (which may be empty). Let \( V' \) denote the one member set containing 1 and let \( V_2 \) denote the set containing \( 2, \ldots, k \).

Now Arrow used the assumption that the individuals can order \( x, y \) and \( z \) in any order and considered a set of individual ordering relations \( R_1, R_2, \ldots, R_n \) such that

1. \( xPy \) and \( yPz \)
2. \( zPx \) and \( xPy \) for all \( i \) in \( V_2 \)
3. \( yPz \) and \( zPx \) for all \( i \) in \( V_3 \)

We notice how Arrow is heading for a problem with transitivity. From 1 and 2 we see that \( xPy \) for all \( i \) in \( V_1 \) but since \( V_1 \) was decisive for \( x \) against \( y \) we conclude that \( xPy \) where \( P \) is the social preference function corresponding to \( R_1, R_2, \ldots, R_n \). From 2 and transitivity we conclude that \( zPy \) for all \( i \) in \( V_2 \) and from 1 and 3 we see that \( yPz \) for all \( i \) not in \( V_2 \).
Now if $zPy$, consequence 2 tells us that $V_2$ is decisive for $z$ against $y$ but this cannot be, because $V_2$ has fewer elements than $V_1$ and we had assumed that $V_1$ was the smallest decisive set for an option over another. Thus we conclude that $yRz$. But since $xPy$ and since $R$ is transitive we conclude that $xPz$.

But from 1 and transitivity we see that $xPz$ and from 2 and 3 we see that $zPx$ for all other individuals than 1. But then we can again use consequence 2 to conclude that the one-member set consisting of individual 1 is decisive for $x$ against $z$. But Arrow had shown in consequence 5 that a one member set could not be decisive for an alternative against any other. “Thus, we have shown that Conditions 1-5 taken together lead to a contradiction” (Arrow 1951, 58).

**Arrow’s Interpretations of the Impossibility theorem**

Arrow formulated this theorem in the following pointed manner:

Possibility Theorem. If there are at least three alternatives among which the members of society are free to order in any way, then every social welfare function satisfying Conditions 2 and 3 and yielding a social ordering satisfying Axiom I and II must be either imposed or dictatorial (Arrow 1950, 342).

He interpreted the result as follows:

If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial (Arrow 1950, 342).

In fact he also emphasized (1951, 31) that, even if one allows addition of individual cardinal utility functions and makes the social choice by maximizing this sum, the method would violate at least one of the conditions set up by Arrow.

Theorem 2 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 discussed in Chapter I Section 1 [Condorcet’s paradox], neither plurality voting nor any scheme of proportional representation, no matter how complicated. Similarly, market mechanism does not create a rational social choice (Arrow 1951, 59).

For this reason Arrow rejected the earlier mathematical treatments of social welfare:

Since, as we have seen, the only purpose of the determination of the maximal states is a preliminary to the study of social welfare functions, the customary study of maximal states under individualistic assumptions is pointless (Arrow, 1951, 63).

**The reception of Arrow’s book and theorem**

In 1972 Arrow was awarded the Nobel Memorial prize in economics. He and John R. Hicks received the prize “for their pioneering contribution to general economic equilibrium theory and welfare theory”. In the press release it was stated:
As perhaps the most important of Arrow’s many contributions to welfare theory appears his “possibility theorem”, according to which it is impossible to construct a social welfare function out of individual preference functions. (Nobelprize.org)\textsuperscript{17}

Arrow’s impossibility theorem is often celebrated as the beginning of modern social choice theory. In the *Handbook of Social Choice and Welfare*, Campbell and Kelly (2002) wrote:

> Arrow’s Theorem … on the aggregation of individual preferences is so startling, and robust, and significant that it spawned a new branch of social studies called social choice theory (Campbell and Kelly 2002, 37).

With even more élan, Amartya Sen (b. 1933), declared Arrow’s impossibility theorem a big bang:

> The big bang that characterized the beginning [of modern social choice theory] took the form of an “impossibility theorem”, viz. Arrow’s (1950, 1951) “General Possibility Theorem”. This theorem… had a profound impact on the way modern social choice theory developed… (Sen 1986, 1074).

In the same vein, one of the main actors in social choice theory Tapas Majumdar (1929–2010) wrote in a review of the book *Algebra of Collective Choice* by another of the key players Amartya Sen:

> So largely do discussions in collective choice theory move round and round Arrow’s inescapable impossibility theorem that often a good first question to ask about any new work on the subject is to call for its bearing in reference to that one fixed point” (Majumdar 1973, 533).

In the introduction to his book *Arrow Impossibility Theorems* Jerry S. Kelly (1978) even likened Arrow’s paper and book to other “extraordinary seminal papers” such as Einstein’s papers on relativity, Gödel’s paper on undecidability, Crick and Watson’s paper on the DNA structure and MacLane’s paper on category theory that all profoundly changed their area of science.

> A similar dramatic case is Kenneth Arrow’s 1950 … paper. This paper (or the book …) has been a principle source for almost all contemporary collective choice theory and has deeply influenced theoretical welfare economics, moral and political philosophy, and mathematical approaches to microeconomic theory (Kelly 1978, 1).

All the above praises of Arrow’s impossibility theorem and his new approach to social choice theory stem from 1972 and later. However, in the early 1950s just after the publication of Arrow’s book, reception was less enthusiastic. Reviewers all emphasized the novelty of Arrow’s mathematical methods, but they disagreed about their usefulness. As usual, non-mathematically inclined economists questioned the necessity of the complicated mathematics:

> When mathematical siege guns are turned on matters previously reserved primarily for verbal discussion, the nonmathematical reader is inclined to inspect the results with an eye biased by his inferiority feelings. What is there here, after all, that we could not have guessed before? Parturient montes, nascitur ridiculous mus.\textsuperscript{18} And so it will presumably be in this case. Yet, however little which

\textsuperscript{17} It is more accurate to talk about individual preference rankings (or orderings) than individual preference functions.

\textsuperscript{18} The mountains are in labor, (and) an absurd mouse will be born. Quotation from Horace.
is substantively new there may be in these conclusions, we must remain grateful to Arrow for caulking the seams of our logic and adding rigor to our intuition (M. Bronfenbrenner 1952, 135).

Most reviewers found the new mathematical formalism difficult:

"Nonmathematical" economists may be prepared to welcome a Cowles Commission monograph that does not use geometry, trigonometry, calculus, difference equations, or determinants. But anyone who expects an hour or two of easy reading is doomed to disappointment. Like most of Von Neumann and Morgenstern’s Theory of Games and Economic Behavior this book employs a particular form of symbolic logic, and the level of concentration required is about the same in both cases (Somers 1952).

But others praised the rigor of Arrow’s approach:

Of interest to many philosophers will be the use which Arrow makes of symbolic logic. Adopting the notation of the logic of relations enables the author to state his conditions in highly perspicuous form and to demonstrate his conclusions with complete rigor (Copi 1952).

At least one far-sighted reviewer sensed that Arrow had introduced the tools of the future:

Perhaps equally interesting is the author’s use of algebraic rather than analytic techniques in his discussion of preference and choice. This approach seems particularly well suited to the subject; it adds to the readability of the volume and lends support to the view that the calculus may not provide the economist his ideal mathematical tools (Baumol 1952).

There were at least five different types of reaction to the impossibility theorem:

1. The pessimistic reaction: The theorem shows that a satisfactory method of voting and social choice is impossible. Democracy is impossible.
2. The theorem is wrong.
3. The theorem is irrelevant to the social sciences.
4. One must relax some of the conditions in order to circumvent the problem proved by the impossibility theorem.
5. One can relax Arrow’s conditions or replace them by other equally (or more) reasonable conditions and still reach an impossibility theorem

Ad 1. In his most pointed formulations of the theorem quoted above Arrow was close to a pessimistic denial of the possibility of a “satisfactory” democratic process. The same holds true of popular formulations intended to bring out the revolutionary and surprising content of the theorem. For example in his presentation speech at the Nobel Award Ceremony the Swedish statistician and national economist Ragnar Bentzel (1919-2005) pronounced:

In his doctoral thesis, which was published in 1951, Arrow put the following question. Let us assume that in a society a number of alternative conditions to choose between and that each individual in the society can rank all these alternatives in order of desirability. Is it, in this case, possible to find ethically acceptable, democratic rules, for making a collective (or social) ranking of the different alternatives in order of desirability? Arrow showed that that question must be answered in the
negative. It is in principle impossible to find such rules. This conclusion which is a rather discouraging one, as regards the dream of a perfect democracy, conflicted with the previously established welfare theory, which had long employed the concept of a social-welfare function. However, this concept is nothing but an expression of a social ranking in order of desirability such as Arrow had shown that it was in principle impossible to make (Nobelprize.org).

However, like most other mathematical impossibility theorems Arrow’s impossibility theorem did not halt development of science nor did it put a stop to democratic ideas or practices.

Ad 2. Most reviews and other comments on Arrow’s paper and book emphasized the mathematical rigor of the argument and the result. However, Julian Blau (b. 1917) (1957) showed that Arrow in his proof of the impossibility theorem had used a stronger version of his conditions than was really warranted from his precise formulation. The problem was that he had used the condition 2 and 5 on the subset of 3 alternatives mentioned in his precise formulation of condition 1, although condition 2 and 5 in the form given to them by Arrow dealt with the whole set of all alternatives. Arrow acknowledged this mistake in the second edition of his book from 1963 (p. 102) and rephrased his conditions accordingly. With this change the impossibility theorem and Arrow’s proof of it have been generally accepted as being mathematically correct.

Ad 3. According to Thomas Kuhn very few scientists are ever converted from one paradigm to another. The supporters of the old paradigm gradually die out and new generations adopt the new paradigm. This pattern fits the Arrowian “revolution” of welfare economics and voting in the sense that the economists of the older generation were not converted. However, it does not fit in another sense; the old paradigm did not die out19. New mathematical economists continued to develop analytic welfare economics of the Samuelson type. So, in this sense Arrow’s contribution cannot be characterized as a Kuhnian revolution.

The older generation accepted the correctness of Arrow’s mathematics or logical argument but the most influential of them declared that the impossibility theorem was irrelevant for welfare economics and for the theory of voting.

a. Black’s reaction: As one would imagine Duncan Black was among the more outspoken critics of Arrow’s result. In the second edition of Social Choice and Individual values, Arrow recognized “the pioneer work of D. Black” (Arrow 1963 (1951 2. Ed), 93). But this did not satisfy Black. In two papers (1969, 1998) he analyzed Arrow’s result as a contribution to his own theory of committee procedures. His argument ran along two tracks. He called attention to his own priority over Arrow for the results he had obtained in particular those concerning single-peaked preferences and he tried to argue that Arrow’s impossibility theorem was irrelevant.

In the first paper, published in 1969, Black showed that unanimity rule satisfied a slight variation of Arrow’s conditions. However, according to Black, this rule is a bad rule because it stalls decisions, whereas better methods such as the Borda count violate the Arrow conditions. From this Black concluded:

(i) A social welfare function is a committee procedure which, when the members’ preferences are subject to certain restrictions, is able to satisfy the Arrow conditions. But since the Arrow conditions

19 Although some proponents of social choice theory have declared the traditional Samuelson approach to welfare economics to be dead (see Igersheim 2017 and Bayjard 2014, 2).
are themselves inappropriate, the social welfare function concerned may be unsuitable as a committee procedure. A social welfare function need not necessarily be chosen in preference to a procedure which is not a social welfare function.

And (ii) the Impossibility Theorem shows that no committee procedure is able to satisfy the Arrow requirements for all sets of preferences. But since these requirements are themselves inappropriate, the fact that no committee procedure is able to satisfy them for all sets of preferences would seem to have no determinate significance (Black 1998, 352).

Black’s negative opinion of Arrow’s results is obvious not only in his argument but also in his rhetoric. This also holds true of the second paper Black wrote about Arrow’s ideas. “Arrow’s work and the normative theory of committees” was published posthumously in 1991, the year Black died. The paper can be considered as Black’s intellectual testament. The essay more directly considered “the relation between the theory of committees” put forward by Professor Kenneth J. Arrow... and ‘the preceding theory’” i.e. Black’s own theories that he also called the positive theory. The paper began with an account of how Black came to his ideas and how he had trouble publishing them. This section that I have drawn on above was clearly meant to establish his priority over Arrow. In the following section, Black used long quotes from his papers of 1949 to show that he had in fact set up norms for voting procedures before Arrow, and he continued to compare Arrow’s norms (conditions) to his own. He argued that two of Arrow’s conditions are more or less equivalent to his own requirements but that a third condition (that of independence of irrelevant alternatives) is misconceived. He maintained that the Arrow conditions cannot be considered as necessary because suitable methods violate them, such as his own preferred method: Majority election if that picks out a Condorcet winner, and if not, a Borda count. The Arrow conditions cannot be sufficient because non-desirable methods, such as unanimity will satisfy them (almost). He suggested “that the omission to consider the extent of intransitivity ... is a great weaknes of the theory” (Black 1998, 397). In particular he referred to his own investigations showing that intransitivity would be an unlikely outcome of many votes and he criticized Arrow’s conditions for not capturing the idea that an election should “chose a motion which can reasonably be held to stand highest on the average” (Black 1998, 392) He called this his statistical thesis.

Referring to “the Kantian principle, which he believed to be generally admitted, that “I ought” implies “I can””, Black wrote of Arrow’s result:

In the context of the search for a suitable committee procedure, holding that we ought to seek to satisfy any condition or set of conditions which it is impossible to fulfill, would be nonsensical (Black 1998, 399).

The only positive thing Black could find to say about Arrow’s work was his novel “treatment in terms of set theory” meaning his use of the axiomatic theory of ordered sets. However, also in this connection Black was critical of Arrow’s use of the word “rational” rather than the neutral word “transitive”, and he pointed out that everything that can be established by way of Arrow’s rigorous axiomatic approach can equally be established in “the previous theory”. Admitting that “by axiomatizing the theory, Arrow’s work had blown a sudden energy into the subject” (Black 1998, 396) he concluded on a more negative note:

To sum up, in accepting the conceptual framework of the preceding theory (i.e. Black’s theory) Arrow deepens one element, replaces another by what would appear to be a wrongly conceived substitute, and yet another element is omitted altogether. He adopts the vision or synoptic view of the earlier theory and, using Symbolic Logic, re-expresses its theory of preference much more effectively. The

---

20 Calling Arrow’s theory a theory of committees was obviously a rhetorical trick that presented Arrow’s contribution as an addition to Black’s own theory bearing that title.
Statistical Thesis of the earlier theory is replaced by the Arrow conditions, which are inappropriate to the task assigned to them; and Arrow’s program of finding social welfare functions can be carried out without advancing the solution of the real problem, which is to find satisfactory committee procedures. Again, to the present writer it seems a mistake to fail to take account of the qualitative aspect in dealing with intransitivities (Black 1998, 403).

Black’s purpose was clearly to put himself forward as the primary inventor of modern theory of elections and depicting Arrow’s contributions as minor technical reformulations and misconceived alterations. From the introduction to the reedition of his works (Black 1998) and the Black Festschrift (Tullock 1991) it appears that some recent economist have indeed considered Black as the father of modern mathematical theory of political economics and social choice.

Coase (1998, xii) emphasized that “it is extraordinary how similar were the paths followed by Black and Arrow” and pointed to one important difference:

While Black worked completely alone, derided and ignored by his colleagues in Britain, Arrow was the recipient of financial support for his research and of moral support from his colleagues at Rand and the Cowles Commission, where he worked with some of the most talented mathematical economists in the world (Coase 1998, xii).

While trying to emphasize the similarity between Black’s (earlier) and Arrow’s (later) accomplishments and relegating the differences to the social domain, Coase completely omitted an evaluation of the essential differences between the two. In particular, by emphasizing the similar results concerning single-peakedness while neglecting the conceptual novelty of Arrow’s mathematical formalism and his impossibility result, Black is made to look more like the true originator of the modern theory than most modern economists are willing to admit. Indeed, while mentioning Tullock’s opinion that Black is “the father of us all” the editors of (Black 1998) admitted that this is “certainly contested” (Black 1998, xxxii). Thus, we see that the discussion of priority that loomed large for both Arrow and Black has been continued by their successors - in particular by the followers of Black who clearly feel that they are up against a majority of supporters of the Nobel Laureate Arrow.

b. The proponents of the analytical “new welfare economics,” Bergson, Samuelson and their follower, the Oxford economist Ian M.D. Little (1918-2012) also renounced the importance of Arrow’s impossibility theorem for their field.\(^2\) The year after the publication of Arrow’s book, Little published an analysis of its methods concluding with the words:

It was first concluded that Arrow’s system is quite different from, and has little relevance to traditional welfare economics. Second, it was found that it cannot without inconsistency be interpreted as a critique of social welfare functions (Little 1952, 432).

Two years later he was seconded by the inventor of the welfare function, Bergson, who agreed “with Little in barring Arrow’s theorem from welfare economics,” concluding:

I have argued that Arrow’s theorem is relevant to political theory, but only to a very special case. The theorem has no bearing on welfare economics (Bergson 1954, 249).

Both Little and Bergson referred to correspondence and conversations with Samuelson, who backed the substance of their criticism publicly in a paper (Samuelson 1967) published four years after Arrow had

\(^2\) Arrow mentions other critics in footnote 32 on p. 103 of the second edition of (Arrow 1951).

Little, Bergson and Samuelson argued that Arrow’s social welfare function was completely different from their function. Therefore the impossibility of an Arrow social welfare function did not imply the impossibility of a Bergson-Samuelson social welfare function as Arrow had implied in his book. In particular they maintained that in welfare economics the tastes of the individuals were fixed, whereas Arrow’s conditions dealt with the effects of variations in values of the individuals. Therefore they denied that Arrow’s conditions were relevant for welfare economics. The subtle differences between the views of Bergson and Little on the one hand and Arrow on the other have been analyzed by Pattanaik (2005), and recently Igersheim (2017) has investigated the controversy between Samuelson and Arrow in the light of unpublished correspondence. Therefore I shall not dwell on this matter.

However, it should be emphasized that welfare economists have continued the approach of Bergson and Samuelson, and in particular the British tradition going back to Kaldor and Hicks after Arrow’s new social choice theory and its new mathematical formalism was embraced by other economists or political theorists. Indeed, it has been argued that the continuation of the older “new welfare economics” has had a greater implication for political practice than the Arrowian mathematically more rigorous social choice theory (Baujard 2014). Thus, instead of considering Arrow’s contribution as a revolution of the field of welfare economics it is more correct to view it as the point of departure of a new revolutionary branch of that wider field.

Ad 4. The most frequent way around the impossibility theorem was to suggest some relaxation of Arrow’s conditions. Arrow himself suggested this way out in his 1950 paper:

If we wish to make social welfare judgments which depend on all individual values, i.e. are not imposed or dictatorial, then we must relax some of the conditions imposed (Arrow 1950, 343).

Maintaining “that there is no meaningful interpersonal comparison of utilities and that the conditions wrapped up in the word “satisfactory” are to be accepted,” Arrow’s only way out was to limit the possible individual orderings:

That is, it must be known in advance that the individual orderings $R_1, \ldots, R_n$ for social actions satisfy certain conditions more restrictive than those hitherto introduced (Arrow 1950, 343).

However, he argued that ethical schemes, such as Bergson’s egalitarian ethics, cannot in themselves rescue the problem raised by the impossibility theorem (Arrow 1951, 72). On the other hand, as we saw above, he had already realized that in the case of single-peaked preference scales, majority voting will lead to a determinate result. At this point he referred to Duncan Black: “A radical restriction on the range of the possible individual orderings has been proposed recently by Professor Duncan Black” (Arrow 1951, 75)\(^{22}\) (see references in Arrow 1951, 6, note 17). He supplied a more rigorous definition of single-peakedness in terms of his axiomatic ordering relations and provided a rigorous proof, that if the domain of the social welfare function is restricted to such orderings, majority voting will satisfy all his conditions (Arrow 1951, 76-80). “An example in which this assumption is satisfied is the party structure of prewar European

\(^{22}\) In footnote 17 on p. 6 of that book, Arrow listed Black’s papers on voting that had appeared up to that time.
parliaments, where there was a universally recognized Left-Right ordering of the parties” (Arrow, 1951, 75-76). Probably because of his independent discovery of the phenomenon and his more rigorous treatment of it, Arrow later referred to single-peakedness as the “Arrow-Black condition” (Arrow and Hervé 1986, 43-46).23

In addition to the possible relaxation of Arrow’s conditions, an assumption of a universally accepted moral or pragmatic imperative, in the sense of Kant or Rousseau, can circumvent the impossibility theorem.

It is overly strong to require that the pragmatic imperatives of different individuals be identical and perhaps even too much to ask that there exist moral imperatives which have this property. The results of Section 2 (single-peaked preferences) show that the condition of unanimity is mathematically unnecessary to the existence of a social welfare function, and we may well hope that there are still other conditions than those laid down there under which the formation of social welfare functions, possibly other than the method of majority decision, will be possible. But it must be demanded that there be some sort of consensus on the ends of society, or no social welfare function can be formed (Arrow 1951, 83).

In the appended notes to the 2nd edition (1963) of Arrows book he defended all his other conditions in the face of suggestions that they were too restrictive and could be weakened. He changed some of the conditions slightly, but the system was essentially equivalent to his previous system.

Yet, in the 1960s and in particular in the 1970s after the appearance of Sen’s book Social Choice and Individual Values many authors suggested relaxations of Arrow’s conditions. In particular the condition of independence of irrelevant alternatives was considered too strong by many authors and so it was relaxed in many ways. Kelly’s book Arrow Impossibility Theorems (1978) gives an idea of the different reformulations of Arrow’s conditions.

Ad 5. In parallel with these attempts to save democratic voting procedures and the welfare function from Arrow’s devastating result, many new impossibility results similar to Arrow’s mushroomed. They all show that a welfare function satisfying a small number of innocent looking and clearly desirable conditions similar to, but not identical to Arrow’s conditions, does not exist. Kelly included about 50 such impossibility results proved by a long list of authors in his monograph (1978) mentioned above, and since 1978 many more have been added to the list (see e.g. Sen 1986). Kelly asserted that the attempts mentioned in the previous section to save the social welfare function by relaxing Arrow’s conditions were hampered by these variants of the impossibility theorem:

The difficulty this reaction (to Arrow’s impossibility theorem) now faces is very simple: for each of Arrow’s conditions, there is now an impossibility theorem not employing that condition (Kelly 1978, 3).

The complementary reactions 4 and 5 mentioned above were responsible for a surge of activity in Arrowian type social choice theory. From about 1970 it marginalized (at least for a period) the “new welfare

23 Coase points to this double naming with implicit disapproval (Coase 1998, xiii).
economics” of the Bergson-Samuelson or the British type and it is now a recognized part of welfare economics with its own journal, society and meetings.

Philosophical impact on mathematical models

Condorcet’s voting procedure and Arrow’s welfare function do the same thing: They transform a set of individual rankings of alternatives (candidates or social states) into a winner or a social choice or a societal ranking. However, the reasons for conducting the vote or the social ranking were very different in the two cases.

For the enlightenment philosopher Condorcet, voting was a method for finding the truth. His favorite example was that of a jury voting on a matter where there was an objective truth (the man killed the victim or he did not). But also in the case of social questions Condorcet believed that there was a scientifically correct answer. The reason for voting was to ensure that individual lack of enlightenment would be almost eliminated from the decision. Similar ideas were expressed by Jean-Jacques Rousseau (1712-1778), Immanuel Kant (1724-1804) and Thomas Hill Green (1836-1882) with various imperatives or a general will replacing Condorcet’s “truth.”

Condorcet considered it a problem if voters let their choices be influenced by personal interests. To him the ideal voter is the enlightened unselfish person:

One cannot consider decisions by majority voting as appropriate for finding out what is true and useful except in the case where a great part of society has the light [is enlightened] & where the people who are educated, who have cultivated their mind, are not subject to prejudices [Condorcet 1785, clxxxii].

Such a view of the voting process was according to Condorcet a result of a historical development. When majority decisions were first used in antiquity they were not designed to arrive at the truth but rather as a method to arrive at “peace and general utility”:

Reflecting on what we know of the constitutions of the ancient peoples, one can see that they sought more to counter-balance the interests and the passions of the different groups (corps) that entered into the constitution of a state rather than obtaining from their decisions results in conformity with the truth (Condorcet 1785, iiij).

However, according to Condorcet in “modern nations” the spread of reason had led to the idea of designing tribunals in such a way that they “would render probable the truth of their decisions” (Condorcet 1785, iij). It is interesting to notice, that, if we accept Condorcet’s historical analysis, ideas oriented around modern welfare are closer to the primitive ancients than to Condorcet’s enlightened view.

Indeed, the ideal “voter” in modern welfare economics is the homo economicus who ranks the social alternatives according to his own interests and opinions. In the original liberal model individuals were supposed to rank alternatives so as to maximize their own material and economic situation. Arrow, on the other hand, explicitly allowed his individuals to make their social choices on the basis of other values than those of the homo economicus, for example “his standards of equity (or perhaps his standards of pecuniary emulation)” (Arrow 1950, 33). However, he emphasized that it is difficult to make a sharp distinction between the tastes of the homo economicus and the values allowed by Arrow, and he maintained that the

---

24 For references see Arrow 1951 p. 81-85.
difference does not change the fundamental problem revealed by the impossibility theorem. And like the utilitarian philosophers and the earlier welfare economists, Arrow rejected the idea that there exists a “true” social state in Condorcet’s sense to which the social welfare function was supposed to find a close approximation:

To assume that the [society’s] ranking does not change with any changes in individual values is to assume, with traditional social philosophy of the Platonic realist variety, that there exists an objective social good defined independently of individual desires. This social good, it was frequently held, could be best apprehended by the methods of philosophic inquiry. Such a philosophy could be and was used to justify government by elite, secular or religious, although the connection is not a necessary one.

To the nominalist temperament of the modern period the assumption of the existence of the social ideal in some Platonic realm of being was meaningless. The utilitarian philosophy of Jeremy Bentham and his followers sought instead to ground the social good on the good of individuals (Arrow 1950, 335).

This is the philosophy behind his condition that the social welfare function cannot be imposed25.

In other words, the aim of the welfare function was not to find the true ideal social state, as had been the purpose of Condorcet’s voting methods, but rather to mediate somehow between the individual interests of the members of society exactly as Condorcet had assumed to be the case in ancient democratic societies. But, despite this great difference in philosophy of voting, the mathematical problems and questions encountered by Condorcet and by Arrow were the same.

This does not mean that philosophical considerations had no influence on the development of social choice theory and its mathematical model. Indeed, we have seen that the rejection, mostly on philosophic grounds, of the possibility of measuring utility cardinally and the rejection of a meaningful intrapersonal comparison of utility, had a crucial impact on Arrow’s new approach. It radically changed the mathematical model of welfare economics from a model based on mathematical analysis (differential calculus) into a discrete or even finite model based on the new axiomatic structure of ordered sets. It was this change that led to the impossibility theorem and created the theory of social choice.

Conclusion. The driving forces

This paper has followed the historical process that led to the creation of a new mathematized branch of the social sciences: Social choice theory. We have seen how it happened through a merger of two hitherto unrelated fields namely welfare economics and the theory of voting. It was Arrow who finalized this merger with his new mathematical model of social choice and his impossibility theorem. Still Bergson and Samuelson had paved the way through their insistence on the ordinal approach to welfare economics and Black had anticipated Arrow by emphasizing the connection between theories of voting and business

---

25 Arrow’s explicit rejection of Platonism in the social realm stands in contrast to his implicit Platonic rhetoric concerning scientific ideas. As we noticed above, when he spoke of his road to the impossibility theorem, Arrow often expressed the belief that the ideas forced themselves on him.
economics (or the theory of the firm). In this section we shall conclude by analyzing the driving forces behind the development.

As pointed out in the previous section, philosophical issues played an important role. Reading Arrow, one is struck by his readiness to engage in philosophical, in particular ethical reflections explicitly referring to philosophers such as Bentham, Kant, Popper and Plato. Ethical considerations underlie the entire question of welfare economics and we have seen how in particular the partly philosophically based distrust in a cardinal measure of utility forced Bergson, Samuelson and in particular Arrow to develop new mathematical models.

Another type of driving forces behind the development could be classified as political. First, the problems raised in the theory of voting as well as in welfare economics is of a political nature. In the opening section of his book, Arrow in particular referred to the “emerging democracies with mixed economic systems, Great Britain, France, and Scandinavia” (Arrow 1951, 1) and the problem of social choice faced in such countries.

Second, the cold war and its institutions played a decisive role in Arrow’s discovery of the impossibility theorem. In particular, his work at the Cowles Commission and the Rand Corporation repeatedly pointed his interest in the direction of the problem of forming some kind of utility function for collectives out of individual utility functions. And when he had made his discovery he was backed by these same institutions and by the influential mathematicians and economists who worked there. In contrast, his competitor Black was working in almost isolation in a country where questions of voting were considered part of recreational mathematics. This difference partly explains his difficulties finding a publisher for his new ideas.

Mixed with this institutional aspect one may sense another political issue namely nationalism. Black felt that he had been treated badly by his American colleagues because they wanted to secure priority for their compatriot Arrow. Indeed, it is conspicuous that the subsequent historical books and papers on Black’s contributions have been written mostly by British authors.

“Priority in discovery is a spur to science” Arrow remarked in connection with his discovery that Black had preceded him in discovering the property of single peaked preferences. At first this discovery had a negative effect on Arrow’s work on voting theories, but it is obvious that his later successes were to some extent motivated by his perceived priority race with Black. In an even more obvious way, Black’s contributions to the theory of committees after 1950 were to a large extent part of his explicit priority dispute with Arrow.

Mathematics provided several driving forces behind the development studied in this paper. First, strive for mathematical rigor was certainly important for Arrow. Indeed, in a period where mathematical economists accused each other of mathematical sloppiness, it was important for the acceptance of Arrow’s new methods and proofs that they were (almost) universally considered to be rigorous. More importantly, it was Arrow’s knowledge of the new axiomatic approach to mathematics, and in particular his knowledge of mathematical logics and abstract ordered sets that allowed him to reformulate the question of social choice and merge it with the problem of voting. It helped him in two ways: First, it provided him with an axiomatic treatment of rankings or orderings enabling him to circumvent Bergson’s and Samuelson’s cardinal indicator. In this way he replaced mathematical analysis with methods of discrete mathematics
which in turn led him to his complete reformulation of the concept of a social welfare function. The axiomatic approach also implicitly helped him in his analysis of the social welfare function or the voting method. Instead of analyzing specific voting methods as earlier voting theorists had done, he proceeded in the spirit of Hilbert’s axiomatic method: he set up a system of axioms or conditions and investigated what one can deduce about a voting method (or a welfare function) that satisfies them, concluding that it does not exist. Without this background in modern axiomatic method neither Samuelson nor Black had been able to make this important step.

As a last driving force one can mention the creative force of impossibility or paradox. Indeed, despite their negative nature, impossibility theorems rarely bar progress. On the contrary they often result in vigorous activity. For example the discovery of incommensurability (the first rigorous impossibility proof) led the ancient Greeks to a great number of important theories. Similarly, the discovery of the unsolvability of the quintic led to the development of Galois theory and much of modern algebra. In the same vein, much of the activity in voting theory was a reaction to Condorcet’s paradox, and the development of social choice theory can be (and has been) considered a response to Arrow’s impossibility theorem. To circumvent the impossible has always been a strong driving force in many areas of life.\(^\text{26}\)

References


\(^{26}\)“Every noble work is at first impossible (Thomas Carlyle)”. This and many other similar quotes collected on the homepage *BrainyQuote* testify to the enduring driving force of the impossible.


Condorcet, Nicolas de (1785) *Essai sur l’application de l’analyse à la probabilités decisions rendues à la pluralité des voix*. L’imprimerie Royale, Paris


Tarski, Alfred (1937) *Einführung in die mathematische Logik und in die Methodologie der Mathematik.* Wien: Springer.

Tarski, Alfred (1941) *Introduction to Logic and to the Methodology of Deductive Sciences*, New York: Oxford University Press.