Rational Analysis of Political Behavior

Social Welfare - Handout

Jorge Fábrega

2024-03-11

Social welfare functions are a key concept in economics and political science, particularly when evaluating the distribution of goods and resources within a society. Nevertheless, its centrality in our analysis will be diminishing across the sessions for reasons that (I hope) will become increasingly obvious sessions after sessions.

For our current interest, welfare functions are used to assess the overall welfare or well-being of a society based on the welfare of its individual members. In this context, we will discuss the definitions and implications of social welfare functions, Pareto solutions, the first and second theorems of welfare economics, the Kaldor-Hicks criterion, and the Arrow impossibility theorem.

Below, you will find a brief summary of each of these topics.

Social Welfare Functions:
A social welfare function (SWF) is a mathematical tool used to rank different social states or allocations based on the utility levels of individuals within the society. Essentially, it aggregates individual preferences into a single measure of societal welfare. The function takes into account the utility/satisfaction/preferences of all individuals and produces a ranking of different societal states from least to most desirable.

There are various forms of social welfare functions, but they all aim to reflect societal preferences in a way that is consistent, non-dictatorial, and respects the Pareto efficiency principle.

Pareto Solution for Social Welfare:
A Pareto solution or Pareto optimality is a state where no individual can be made better off without making at least one individual worse off. In the context of social welfare, a Pareto optimal allocation is one where the welfare of society cannot be improved without reducing the welfare of at least one individual. It is a baseline criterion for efficiency in economics, indicating that all possible mutual benefits from trade or reallocation have been achieved.

First Theorem of Welfare Economics:
The first theorem of welfare economics states that any competitive equilibrium leads to a Pareto optimal allocation of resources. This implies that under perfect competition, where there are no externalities, public goods, or market imperfections, the market will naturally allocate resources in a way that maximizes societal welfare. The theorem highlights the efficiency of market systems under certain conditions, suggesting that in an idealized market environment, individual pursuits of self-interest will lead to socially optimal outcomes.

Second Theorem of Welfare Economics:
The second theorem of welfare economics states that any Pareto optimal allocation can be achieved as a competitive equilibrium, given appropriate redistribution of initial endowments. This implies that, with suitable adjustments to the initial distribution of resources, any desired equitable outcome can be reached without losing efficiency. The theorem underscores the separability of efficiency and equity concerns, suggesting that redistribution policies can be used to achieve desired social welfare outcomes without compromising the efficiency benefits of market systems.

Kaldor-Hicks Criterion:
The Kaldor-Hicks criterion is a compensation test used to evaluate economic improvements or policy changes.

It states that an action is considered to improve societal welfare if those who are made better off *could hypothetically compensate* those who are made worse off, and still be better off themselves. Unlike the Pareto criterion, the Kaldor-Hicks criterion does not require that no one be made worse off—only that the potential gains outweigh the losses. It allows for the assessment of policies that produce winners and losers, as long as the net effect is positive.

Example: The government plans to construct a new highway to reduce traffic congestion, which will benefit commuters (Group A) by decreasing travel time. However, the construction will lead to the demolition of a public park, negatively impacting local residents (Group B) who use the park for recreation.

Utility Expressions:
Utility of Group A (commuters): $U_A = V_A - C_A$, where V_A is the value of time saved and C_A is the cost shared for the highway construction.
Utility of Group B (local residents): $U_B = V_B - L_B$ where V_B is the value of the park to residents and L_B is the loss from the park's demolition.

According with the K-H criterion: Should the project be executed is $V_A = 500$, $C_A = 200$, $V_B = 300$, and $L_B = V_B$ (total loss since the park is demolished)?

Any thoughts?

Cardinality versus Ordinality
The Kaldor-Hicks (K-H) criterion, unlike the Pareto Efficiency Criterion, allows for compensation between individuals, effectively permitting policies that may harm some while benefiting others, as long as those who benefit could *in theory* compensate those who lose. This approach adopts a cardinal utility perspective, implying that individual utilities can be compared and aggregated, which is a significant departure from the ordinal nature of Pareto efficiency that avoids such comparisons.

The implicit assumption underlying the K-H criterion aligns with utilitarian social welfare functions, where the total or average utility is maximized regardless of its distribution among individuals. This raises ethical and practical concerns.

By relying on the K-H criterion, one implicitly assumes a specific type of social welfare function that may not respect all individuals' welfare equally, potentially violating the non-dictatorial and consistency conditions of a fair and equitable social choice mechanism. Furthermore, the application of the K-H criterion can conflict with the Pareto principle when compensations are theoretical rather than actual, leading to situations where policies are deemed acceptable under K-H but result in real-world Pareto inefficiencies. This highlights the complex trade-offs between efficiency, equity, and procedural fairness in economic policy-making and the inherent challenges in constructing a social welfare function that simultaneously meets all desirable criteria.

To look at this problem. Let consider the role of Public Goods and Externalities.

Example 1: Public Goods and Climate Change Mitigation
Scenario: A government is considering the installation of a new large-scale solar power plant to reduce carbon emissions. This is a classic public good since its benefits (reduced emissions) are non-excludable and non-rivalrous. Hence, everyone benefits from a cleaner environment and the mitigation of climate change, regardless of whether they contribute to the funding of the solar plant.

Kaldor-Hicks Criterion Issue: According to the Kaldor-Hicks efficiency, the project should proceed if the total economic benefits exceed the costs, even if the losers are not actually compensated. In this case, the direct beneficiaries might be the local communities with cleaner air and renewable energy. However, the cost might be borne by the entire nation through increased taxes or by specific groups, such as workers in the fossil fuel industry who may lose their jobs, or landowners where the solar plant is built who may suffer from reduced property values.

Practical Problem: While the project may lead to an overall increase in social welfare (reduced carbon emissions contributing to global climate mitigation), the Kaldor-Hicks criterion doesn't address the distributional effects or ensure that the parties who incur the costs (like displaced workers or impacted landowners) are compensated. This can lead to resistance from affected groups, political backlash, and potential social inequity, undermining the sustainability and public acceptance of climate change mitigation efforts.

Example 2: Externalities and Adaptation to Climate Change
Scenario: A coastal city decides to build a seawall to protect against rising sea levels and storm surges, a response to climate change. This infrastructure project will benefit property owners and residents in flood-prone areas by reducing the risk of water damage and loss of property.

Kaldor-Hicks Criterion Issue: Under the Kaldor-Hicks framework, the construction is justified if the aggregate benefits (prevention of flood damage) exceed the aggregate costs (construction and maintenance of the seawall). However, this calculation may not account for the negative externalities, such as environmental damage to marine ecosystems or reduced access to the beach for the public. Additionally, the financing of the seawall might come from general municipal funds, affecting taxpayers who live in higher elevations and are not at risk from flooding.

Practical Problem: The implementation based on the Kaldor-Hicks criterion might lead to environmental degradation and social issues, as the interests of minority or less influential groups (e.g., environmentalists, non-coastal residents) are overlooked. Moreover, the criterion does not ensure that those who bear the costs (taxpayers, ecosystem services users) are appropriately compensated or that the benefits are equitably distributed, especially if wealthier property owners receive disproportionate protection compared to more vulnerable populations.

Looking at the Pareto versus Kaldor-Hicks criteria using Edgeworth Boxes

Edgeworth boxes are an essential tool in understanding the dynamics of public policymaking and the distribution of social welfare. This section will guide you through constructing Edgeworth boxes and interpreting their implications for identifying winners and losers in public policy changes.

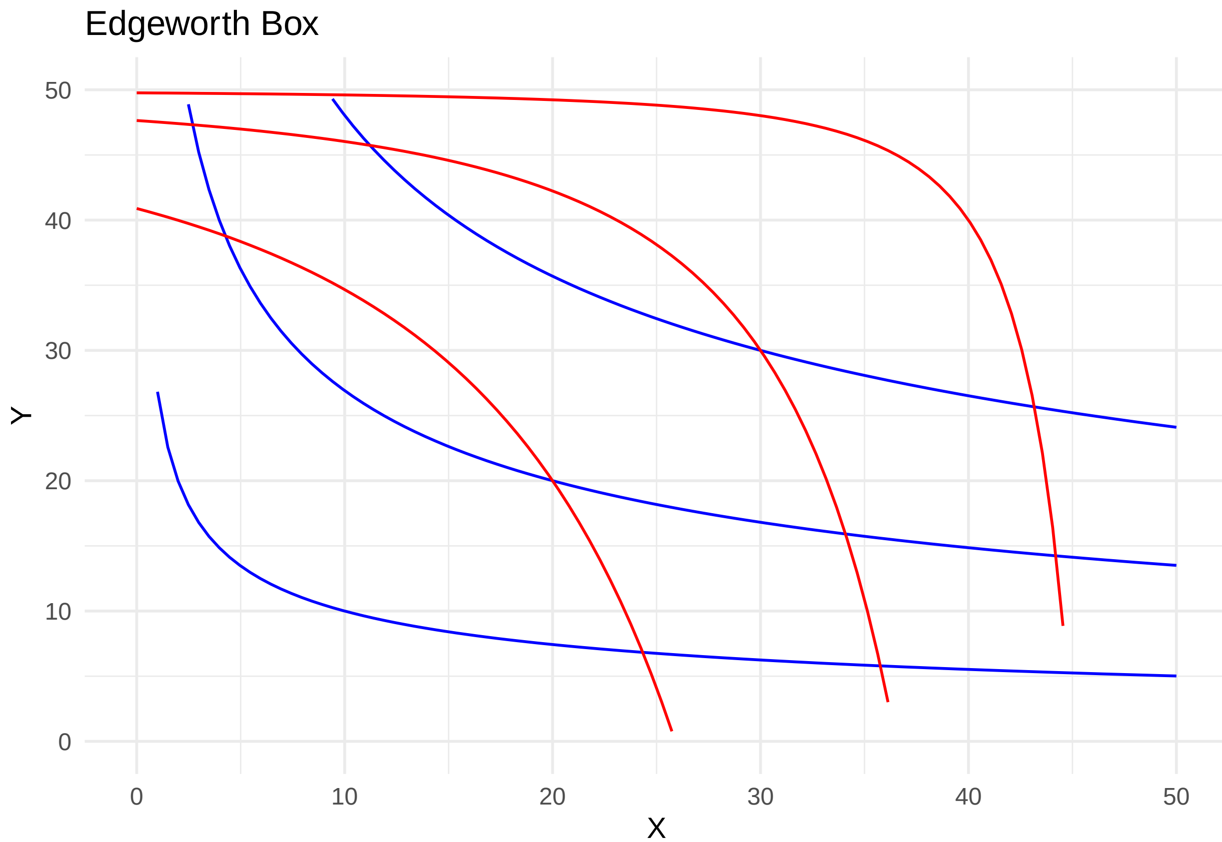
The Edgeworth box is a graphical representation derived from the concepts of microeconomics, specifically the theory of exchange. It illustrates how different allocations of resources can lead to changes in social welfare and individual utility.

Constructing Edgeworth Boxes:

To build an Edgeworth box, we begin with two individuals, A and B, and two goods, X and Y. Each axis represents the total amount of one good, with the origin for individual A at the bottom left of the diagram and the origin for individual B at the top right. This dual-origin setup allows us to visualize each individual's consumption of goods X and Y simultaneously.

The length and width of the box represent the total supply of goods X and Y, respectively. Inside the box, we draw indifference curves for each individual. These curves represent combinations of goods X and Y that provide the same utility or satisfaction level to the individual. For A, the curves bow towards B's origin, and for B, they bow towards A's origin. The area where these indifference curves intersect is known as the **contract curve**, which represents all the potential Pareto-efficient allocations—situations where no one can be made better off without making someone else worse off.

The following figure represents an Edgeworth box and two sets of indifference curves:



Interpreting Edgeworth Boxes:

The interpretation of Edgeworth boxes is central to understanding the impacts of public policymaking. Consider a policy change aimed at reallocating resources between two sectors in an economy, represented by goods X and Y. By observing how the policy shifts the allocation of goods within the Edgeworth box, we can identify who gains or loses from the policy.

For instance, if a new policy provides subsidies for good X, which is primarily consumed by individual A, we would see A's indifference curves shift towards more favorable allocations of X. As a result, A might move to a higher indifference curve, indicating an increase in welfare. Conversely, if B relies more on good Y and the subsidy leads to a relative increase in the price of Y, B's welfare may decrease as they move to a lower indifference curve.

But there are also theoretical concerns, especially when considering the impossibility theorems in welfare economics, such as Arrow's Impossibility Theorem.

Examples and Applications:

To illustrate, consider a government deciding to subsidize education (good X) over military spending (good Y). In an Edgeworth box, we could represent the total resources dedicated to education and military on the two axes. Citizens favoring more education would move towards higher indifference curves with the subsidy, reflecting their increased welfare. In contrast, those prioritizing military spending would experience a decrease in utility, moving to lower indifference curves due to reduced allocation to good Y.

Another example. Suppose there is a city divided into two areas: Eastside and Westside. Both areas have the same population size, but they start with different endowments of healthcare resources and wealth. Health reform aims to redistribute these resources more equitably between the two areas.

Data Provided:

Total healthcare resources (H) available in the city: 100 units.

Total wealth (W) available in the city: 200 units.
Initial endowment in Eastside: 70 units of healthcare and 80 units of wealth.
Initial endowment in Westside: 30 units of healthcare and 120 units of wealth.

Instructions:

Setting up the Edgeworth Box:
Draw the Edgeworth box. The length of the box should represent the total healthcare resources (100 units), and the width should represent the total wealth (200 units). Plot the initial endowment point for each area based on the data provided (Eastside: 70H, 80W; Westside: 30H, 120W).

Drawing Indifference Curves:
Draw a set of indifference curves for each area starting from their respective initial endowments. Remember, the curves should reflect each area's preferences for healthcare and wealth. Indifference curves closer to an area's origin represent lower levels of satisfaction.

Identifying the Contract Curve:
Identify and draw the contract curve within the box. This curve should connect all points where Eastside's and Westside's indifference curves are tangent to each other, indicating all efficient allocations of healthcare and wealth.

Analyzing Health Reform:
Assume the health reform aims to redistribute resources to achieve a more equitable distribution. Discuss with your classmates what an equitable distribution might look like and mark this new allocation on the Edgeworth box.

Shift the allocation point from the initial endowments to the new equitable distribution you have identified. Observe how this shift moves along or towards the contract curve. Discussion Questions:

How do the initial and new allocations compare in terms of equity and efficiency?
Which area gains and which loses from the health reform? By how much?
Given those gains and loses, should the reform be implemented according with the K-H criterion?
Is there a way to make both areas better off (Pareto improvement)? Discuss potential trade-offs.
Reflect on the real-world implications of this analysis for health policy.

Now, let consider one more time these trade-offs under the lens of cardinality versus ordinality. What are we assuming on the comparison between East and West side utilities under K-H criterion?

Given the above scenario, we should move to a theoretical caveat on KH criterion and the (unsuceesful) search for a solution. The guiding question is what the required principles to build a social welfare function are. Any idea?

Arrow's impossibility theorem
The Arrow's Impossibility Theorem, formulated by economist Kenneth Arrow in the early 1950s, is a fundamental result in social choice theory and welfare economics. The theorem, also known as Arrow's Paradox, states that no rank-order voting system can meet all of the following fairness plus consistency criteria when there are three or more options (or candidates):

Non-dictatorship: The preferences of an individual should not become the group's ranking without considering the others' preferences. That is, there is no single voter whose preferences dictate the group's preferences.

Unrestricted Domain (or universality): The voting system should account for all possible individual preferences among the options.

Independence of Irrelevant Alternatives (IIA): If one option is removed from consideration, the collective preference between the remaining options should not change.

Pareto Efficiency: If every individual prefers one option over another, then the group should also prefer that option over the other.

Transitivity: If the group prefers option A over B and B over C, then the group should prefer A over C. Arrow demonstrated that no voting system could satisfy all these conditions simultaneously if the choice set contains three or more options. This result is surprising and powerful because it shows that all voting systems have inherent limitations.

Examples of Arrow's Impossibility Theorem:

Majority Voting (First-past-the-post):

Non-dictatorship: This condition is met since no single voter controls the outcome.
Unrestricted Domain: This condition is met as voters can rank candidates in any order.
Pareto Efficiency: Generally met if all voters prefer one candidate over another.
Transitivity: In a direct head-to-head comparison, this is usually met.
IIA: This is where majority voting fails. Consider three candidates: A, B, and C, with the following preference profile among three voters:

Voter 1: $A > B > C$
Voter 2: $B > C > A$
Voter 3: $C > A > B$

In this scenario, there is a problem with the transitivity requirement. Can you see, why?.

Ranked Choice Voting (Instant Runoff): Non-dictatorship: This condition is met since no single voter's preference dictates the outcome.
Unrestricted Domain: This condition is met as voters can rank their preferences freely.
Pareto Efficiency: This condition is generally met within this system.
Transitivity: This condition is met within individual voters' rankings.
IIA: Again, this system fails the IIA condition. Suppose a voter ranking as follows among four voters:

Voter 1: $A > B > C > D$
Voter 2: $B > A > C > D$
Voter 3: $C > D > A > B$
Voter 4: $B > C > D > A$
Voter 5: $A > C > B > D$ Voter 6: $C > D > A > B$ Voter 7: $D > A > B > C$

Initially, no candidate has a majority. If D, the least popular, is eliminated, their votes might transfer to A, changing the overall outcome. Such that in a second round, A gets 3 votes, and B and C only 2 votes. The next to be eliminated will be between the last two. Consequently, we should be ready to state that A is more desirable (socially speaking) that B or C.

But if a new candidate, E, is added, the elimination order could change, affecting who gets transferred votes and ultimately who wins, thus violating the IIA condition. For example:

Voter 1: $E > A > B > C > D$
Voter 2: $B > A > C > D > E$
Voter 3: $C > D > A > B > E$
Voter 4: $B > C > D > A > E$
Voter 5: $E > A > C > B > D$ Voter 6: $C > D > A > B > E$ Voter 7: $D > A > B > C > E$

In this case, the first to be eliminated is A. Consequently, the introduction of a new (supposedly irrelevant) alternative changed the social desirability of A with respect to B and C.

Kenneth Arrow proved that this impossibility holds for any imaginable voting system with at least three options and voters.