

Rational Choice, Utility Functions, Welfare Functions, and Policymaking

In an article discussing various definitions of economics, Backhouse and Medema (p.225, 2009) present Lionel Robbins's definition of economics: "the science which studies human behavior as a relationship between ends and scarce means which have alternative uses," and remark that this particular definition is "perhaps the most common currently accepted definition of economics."

Human behavior, however, is notoriously hard to model. People are diverse and their decisions are influenced by a multitude of ever-changing factors: monetary incentives, expectations and beliefs, emotions, morals, cultural influences, social pressures, forces of habit, etc. This complexity is further compounded once we take into consideration direct interactions between humans as well as indirect interactions among humans through institutions such as markets. Due to the overwhelming complexity involved in modeling human behavior, economists typically make simplifying assumptions to create models that can be analyzed and solved.

These simplifying assumptions cause economic models to sometimes appear to be unrealistic and wrong. However, that should not lead us to dismissing economic models. As the famous statistician George Box wrote, "All models are wrong, but some are useful" (Box 1978, 202). A common analogy compares models to maps. A map showing how postal codes or zip codes are organized within a city will be very different from a map showing the subway routes in the city. While both maps reflect aspects of reality pertinent for their own purpose, they are simplified in that they don't include details that are unnecessary for their purpose.

While there are multiple approaches to modeling human behavior within the discipline of economics, the most common approach assumes that people are capable of **rational choice**.

I. Rational Choice Models

In the rational choice framework, people are modeled as goal-directed '**utility**' or happiness maximizers. In other words, economists assume that people make choices that are optimal in terms of enabling them to become happier. For example, a person who chooses to watch a horror movie does so because the excitement that comes from being frightened or horrified while watching the movie ultimately makes that person happier than any other alternative activity available at the time. Similarly, when a person decides to donate some money to a charitable cause it is because giving money to the charity makes the person happier than any other alternative use of the money.

Within the framework of rational choice, people are also assumed to be perfectly well-informed about all their choices as well as perfectly able to calculate their optimal choice in all situations. Since these assumptions are somewhat unrealistic, economists refer to the people within their idealized models as belonging to the fictitious species, *homo economicus*. Unlike *homo sapiens*, members of *homo economicus* always act rationally and are perfectly informed and unconstrained by limitations of memory or computing power.

Only four simple assumptions are needed for rational choice models, two of which are mathematical while the other two are conceptual:

1. Information: People know about all the choices available to them as well as the outcome associated with each choice.

2. Completeness: People can make pairwise comparisons across all outcomes and recognize whether they prefer one or the other outcome or are indifferent between the two.
3. Transitivity: If an agent prefers outcome 'a' to 'b' and also prefers 'b' to 'c,' she necessarily prefers outcome 'a' to outcome 'c.' Together with assumptions 1. and 2., this implies agents can consistently and reliably rank all their choices from most to least preferred.
4. Optimization: Agents choose their most preferred outcome.

When combined with the idea that a person's tacit preferences are revealed by their actual choices, preferences can be expressed as fully specified mathematical **utility functions** that can be estimated from data on people's choices. Utility functions enable economists to conduct mathematical analysis very conveniently (e.g., they can use calculus to find optimal decisions). With the inclusion of two further technical assumptions (continuity and independence), rational choice models can be extended to incorporate uncertainty.

While there are many serious criticisms of rational choice models, it is an approach that has allowed economists to conduct rigorous mathematical analysis of complex problems where people (buyers, sellers, workers, farmers, financiers, etc.) must make choices that inevitably involve tradeoffs.

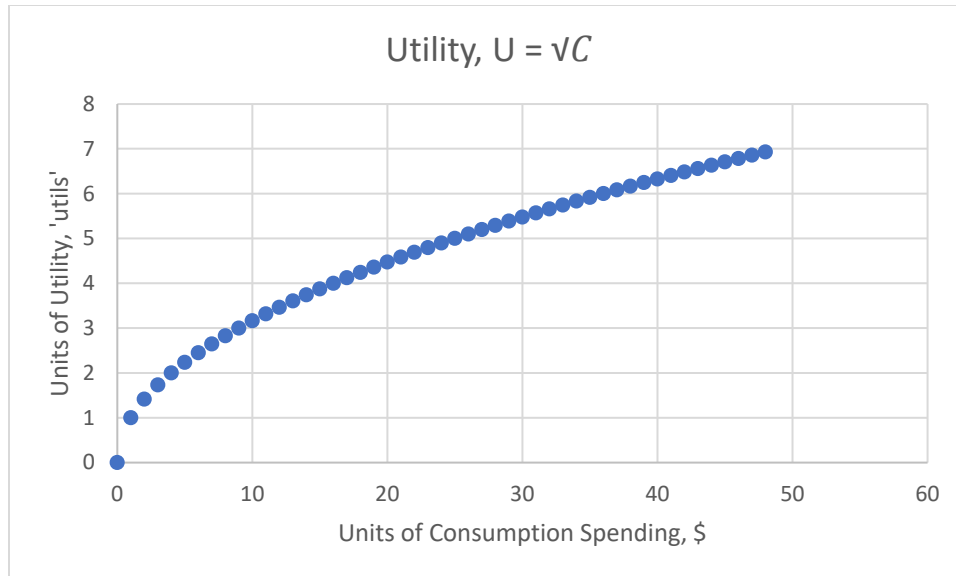
II. Utility Functions

A utility function is a mathematical function that specifies the level of happiness a person will experience based on their economic circumstances. While a person's utility or happiness is not directly measurable and is inherently subjective (what gives one person happiness might differ significantly from what gives another person happiness), economists try to make general estimates of these functions by studying large datasets of people's actual choices.

Depending on the specifics of the situation being modeled, an economist may assume that a person's utility depends on how much they are able to spend on consumption goods, how much leisure time they are able to enjoy, the state of their health, etc. A very simple utility function, used primarily for illustrative purposes in high school and undergraduate economics courses, is the following:

$$U = \sqrt{C}$$

According to this simple utility function, a person's happiness can be measured as the square root of their level of consumption spending. This utility function incorporates the idea of **diminishing marginal utility** – a psychologically intuitive idea that is also robustly borne out by data on people's actual behavior. As can be seen from the graph below, diminishing marginal utility is the idea that at low levels of consumption an extra unit of consumption can add a great deal to happiness whereas at high levels of consumption, an extra unit of consumption does not add much to happiness. Mathematically, when consumption increases from 0 to 1, utils also increase by 1. However, when consumption increases by 1 unit from 24 to 25, utils increase by only 0.1. In the context of food, the first slice of pizza can bring a great deal of happiness to a hungry teenager but the fifth or sixth slice will bring much less happiness. The simple idea of diminishing marginal utility can also be used to explain **risk aversion** (people buy insurance to protect against potential losses) and **inter-temporal consumption smoothing** (people save during periods when their incomes are high so that they can finance their living during periods when their incomes are low).



Depending on the problem being studied, economists could include several other features into mathematical utility functions. I am listing a few below:

- Preference for variety – some utility functions incorporate the idea that a person’s level of happiness increases when they are able to consume a greater variety of products. In the context of food, this could mean that a person gets more happiness from a combination of coffee and cupcakes than from coffee alone or cupcakes alone.
- Preference for quality – some utility functions incorporate the idea that people recognize quality differences between different types of consumption. For instance, based on data on consumption patterns, economists know that when people become richer they tend to increase their consumption of meat (or other types of protein for those who are vegetarian for religious or cultural reasons) while decreasing their consumption of potatoes. This suggests that people, in general, think of meat as higher quality food than potatoes.
- Altruism or benevolence – some utility functions incorporate the idea that people derive happiness from the wellbeing of others – this could explain why people tend to donate to charities, help out their friends in times of need, etc. Similarly, sometimes economists assume that utility functions include some component of social responsibility.
- Loss aversion – some utility functions include the concept of loss aversion. According to the concept of loss aversion, when a person gets habituated to a certain style of consumption, changing their style of consumption causes greater disutility than if their consumption style were to stay the same. As an example, it is harder for someone who is accustomed to watching Netflix to give it up than it would be for someone who does not watch Netflix to continue without Netflix.

III. Social Wellbeing

Within economic models, utility functions can also indirectly help to measure some overall level of social wellbeing or social welfare. While an individual's utility can be mathematically defined as a function of their individual level of consumption, their individual level of leisure time, their individual level of health, etc., a society's overall wellbeing can be defined in terms of the utility levels of all the individuals in the society. It might best to demonstrate this with a simple example:

Assume a society with 5 individuals where the first person has a consumption level of 1, the second has a consumption level of 25, the third has a consumption level of 16, the fourth has a consumption level of 100, and the fifth has a consumption level of 49. In mathematical notation: $C_1 = 1$, $C_2 = 25$, $C_3 = 16$, $C_4 = 100$, and $C_5 = 49$. Assuming the simple utility function where $U = \sqrt{C}$, their corresponding utility levels are: $U_1 = 1$, $U_2 = 5$, $U_3 = 4$, $U_4 = 10$, and $U_5 = 7$.

While there are innumerable many ways of defining social wellbeing or social welfare, here I will discuss three common approaches with economics.

1. **Standard or Average Welfare Function:** This is simply the average of the utility values. $\frac{U_1 + U_2 + U_3 + U_4 + U_5}{5} = \frac{27}{5} = 5.4$. According to this measure of social wellbeing, the happiness level of this society is 5.4 and any policy that increases this happiness level to a higher number would be preferable to the current situation. Given a fixed population size, it is often easier to use the sum rather than the average since the relative ranking of averages across the various scenarios under consideration will be the same as the relative ranking of sums across those scenarios.

A common variant of the standard or average welfare function involves a *weighted* average of utilities rather than a simple average of utilities. These weights could differ across individuals based on several different criteria such as the time-period under consideration (with utilities accruing in the future 'discounted' by time¹), the citizenship status of the person under consideration, the political influence of the person under consideration, etc.

2. **Rawlsian Welfare Function:** According to the philosopher John Rawls, society should primarily focus on making whoever is the worst-off better to the extent possible.² Several religious thinkers and philosophers also emphasize that a society should be judged by how it treats its most vulnerable³. Based on this idea, the overall social wellbeing of this society is given by the value of utility for the person with the lowest utility. Mathematically, according to the Rawlsian welfare function, social wellbeing for this society is given by $\text{Min}(U_1, U_2, U_3, U_4, U_5) = 1$. In the example provided, any policy that increases the happiness level of the first person to a value greater than 1 (without lowering the happiness level of any other individual to a value less than 1) would make society better off.

¹ The appropriate discount rate for social welfare functions is a topic of intense debate among economists and policymakers.

² According to Rawls, the first and foremost requirement is that all individuals need to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all. The concern for the 'worst off' is of secondary importance.

³ Opinions might differ on who is considered the most vulnerable. Some might argue that a refugee fleeing a brutal civil war is the most vulnerable while another might consider an unborn fetus to be the most vulnerable.

3. **Kaldor-Hicks Efficiency (sometimes referred to simply as efficiency):** This approach avoids utility functions altogether and is the most common approach found in high school and college economics textbooks. This measure of social wellbeing focuses on the total amount of economic resources available in society. In the context of the simple example being discussed here, it would simply be the sum of all the values of consumption across all individuals. In more complicated models this could equal the sum of producer surplus, consumer surplus, and government revenue, or it could equal GDP, or it could equal some complex combination of consumption, leisure time, etc. Mathematically, for the simple example being discussed here, social wellbeing can be given by $C_1 + C_2 + C_3 + C_4 + C_5 = 191$. Any policy that increases overall consumption to a value greater than 191 would therefore be a good policy. Economists who use this policy defend it by arguing that an economist's focus should be to increase overall economic resources while some other political process should be invoked for the subsequent decision of how these resources should be distributed across people in society.⁴ Another pragmatic argument for Kaldor-Hicks efficiency is that it avoids interpersonal utility comparisons. The two welfare functions discussed previously implicitly compare across the utility values of different people. However, it is certainly possible that different people obtain utility in vastly different ways and comparing across them might therefore be theoretically impossible.

IV. Positive and Normative Economics and their Relationship with Policymaking

A positive statement is a *descriptive* statement. While typically factual, it could also include contentious (or even incorrect!) opinions. Positive economics is ultimately merely a description of the real economy or an economic model. A statement such as "in my view, the unemployment rate was 35% during the year 2016" is a positive statement though it is a factually incorrect opinion. A statement such as "raising the minimum wage will increase unemployment" is also a positive statement (and an opinion) though it is somewhat contentious.

A normative statement is a *prescriptive* statement. It involves an implicit (or explicit) value system that ranks some outcomes as being preferable to other outcomes.⁵ An example of a normative statement could be, "the Federal Reserve should focus more on unemployment than inflation."

Policy recommendations typically involve positive and normative components. As an example, consider the following policy recommendation: "the Federal Reserve should lower the interest rate to fight rising unemployment." Like most policy recommendations in economics, this statement also has some kind of positive economics as the first part of its foundation and some kind of normative value system as the second part of its foundation. In the example above there are at least two positive ideas: 1) unemployment is rising, and 2)

⁴ According to Kenneth Arrow separating the distributional concerns from efficiency concerns is theoretically possible only if certain stringent assumptions required for two technical theorems (commonly known as the 1st and 2nd welfare theorems) are satisfied. Arrow (1963, p. 943): "If, on the contrary, the actual market differs significantly from the competitive model, or if the assumptions of the two optimality theorems are not fulfilled, the separation of allocative and distributional procedures becomes, in most cases, impossible."

⁵ It may be interesting to note that from the perspective of economic policy makers, utility functions are often considered to be positive features, i.e., the policymaker typically assumes that individuals' utility functions are what they are given assumptions about human psychology. However, from the perspective of parents, pastors, philosophers, teachers, and individuals themselves, utility functions could also be considered normative because parents, pastors, philosophers, teachers, and individuals may have a preference ordering among the utility functions themselves e.g., they might want utility functions to be more prosocial, more pragmatic, etc.

lowering interest rates can help reduce rising unemployment. In addition, there is the normative value system that prioritizes fighting unemployment as the appropriate goal for the Federal Reserve.

When economists disagree on policy prescriptions, they may be doing so for either or both of the following two reasons:

1. They disagree about the positive questions (e.g., what the situation in the economy is, which economic model is appropriate to analyze the situation, which dataset is the closest fit for quantitative guidance, which method of analysis is the most suitable, etc.).
2. They have different normative value systems. Therefore, even if they agree on the outcome of the proposed policy, they may disagree on whether it is the right policy (e.g., one economist might value growth while the other might value equality, one economist might argue that inflation is worse than unemployment, while the other might argue the opposite, etc.).

V. Criteria to Evaluate Economic Policy

Economists often use models to analyze policy recommendations. Their approach is therefore inherently **consequentialist**, i.e., their recommendations are based on the projected consequences of the various policies under consideration. Even within the consequentialist approach, depending on the model or policy being considered, recommendations can be based on either **proximate considerations** or **ultimate considerations**.

- Proximate considerations often involve specific objectives such as lower poverty, lower inflation, lower unemployment, higher growth, lower inequality, etc. In many situations, there are tradeoffs among these objectives and different normative value systems weigh these different objectives differently. These considerations are often directly measurable metrics and hence convenient for expedient analyses.
- Ultimate considerations, on the other hand, involve a more foundational philosophical framework and require explicitly recognizing why and how proximate considerations such as inflation, unemployment, growth, equality, etc. ultimately affect people's utility levels and therefore overall social wellbeing. In terms of ultimate considerations, policies can be evaluated using several different frameworks among which are three measures of social wellbeing discussed above.
 - Standard or average welfare – according to this criterion, a policy is desirable if it increases social wellbeing as measured by the standard or average welfare function.
 - Rawlsian welfare – according to this criterion, a policy is desirable if it increases social wellbeing as measured by the Rawlsian welfare function.
 - Kaldor-Hicks efficiency – according to this criterion, a policy is desirable if it increases social wellbeing as measured in terms of Kaldor-Hicks efficiency.
- Policies can also be judged in terms of another criterion known as **Pareto improvement**. A policy results in a Pareto improvement if it makes at least one individual better-off than they currently are without making anyone worse-off than they currently are. This criterion also has the advantage of avoiding direct interpersonal utility comparisons because all that needs to be known is whether anyone's utility level would go down for a policy to be rejected. While very appealing since no one would argue against a Pareto improving policy, this criterion is hard to apply in many real-world policy problems because most economic policy changes result in some people becoming better-off while others become worse-off. Moreover, this evaluation criterion results in a **status quo bias**. For example, if the government currently has a policy of free

college, then removing the policy would decrease the utility of students so this criterion would forbid removing the policy; on the other hand, if the government currently has no policy of free college, then introducing such a policy would also be forbidden because it would decrease the utility of some taxpayers who ultimately pay for the free college. Similarly, if there is no minimum wage regulation, the implementation of such a regulation would violate Pareto improvement since the regulation would hurt at least some people in the economy, for instance some employers who would have to pay higher wages. On the other hand, if minimum wage regulations are already in place, removing them would also violate Pareto improvement since at some people in the economy would be hurt by the change, for instance some of the workers who currently earn the minimum wage would receive lower wages. When an economy achieves a situation such that no further Pareto improvements are possible, it is said to be **Pareto efficient** or **Pareto optimal**. Since the concept of Pareto efficiency or Pareto optimality is of great importance in economic theory, it might be useful to discuss some aspects related to it.

- If we assume that people are primarily selfish (or concerned primarily about themselves), then Pareto improvement and even Pareto efficiency can be determined without making interpersonal utility comparisons. As discussed previously, some economists and philosophers find this feature to be appealing because different people might obtain utility in vastly different ways and comparing across them would therefore be theoretically impossible.
- If an outcome is **not** Pareto efficient, by definition, at least one individual can be made better off without making anyone else worse off. Based on this reasoning, many economists argue that all desirable final outcomes need to be Pareto efficient. However, most economists would also agree that not all Pareto efficient outcomes are desirable. For example, an outcome where one selfish individual owns everything while others have nothing is Pareto efficient (since no one can be made better off without hurting the rich individual) but would be considered unfair by most.
- It can be shown mathematically that idealized free markets (i.e., perfectly competitive and complete markets with no externalities, no asymmetric information, etc.) will necessarily lead to a Pareto efficient outcome⁶ and will do so entirely through Pareto improving voluntary transactions. However, the assumptions required for idealized free markets are often violated; many real-world markets suffer from issues such as externalities, asymmetric information, overwhelming market power for some participants, etc.
- *Every* Pareto efficient outcome can necessarily be expressed mathematically as the optimal outcome for a corresponding *weighted* standard welfare function. It might be of interest to note that the weights in the weighted social welfare function associated with the Pareto optimal outcome resulting from free markets might be extremely unequal across individuals.⁷ In other words, free markets do not address inequality and could even exacerbate it. To get an intuitive sense of why this is so, recall that every transaction in free markets needs to be Pareto improving. Therefore, even

⁶ This is known as the 1st welfare theorem (this was also alluded to in footnote #3).

⁷ These weights are called Negishi weights. Mathematically, it can be shown in simple models that the Negishi weight for an individual is inversely proportional to that individual's marginal utility in the Pareto efficient outcome achieved through free markets. Deriving this is straightforward but beyond the scope of this handout.

those whose initial utility (or consumption) levels are very high will necessarily see further increases in their utility levels.

- **Kaldor-Hicks efficient** outcomes (i.e., outcomes that maximize Kaldor-Hicks efficiency) are sometimes referred to as a **Potentially Pareto efficient** outcomes. This is based on the reasoning that if a certain policy causes the overall economic pie to grow to its maximum while possibly hurting some individuals (i.e., the policy could violate the criterion of Pareto improvement), some *potential* theoretical costless redistribution (not considered to be part of the policy) from the winners to the losers could improve material outcomes and therefore wellbeing for everyone in the economy resulting in an unambiguous Pareto improvement with no further Pareto improvements impossible.
- Amartya Sen championed yet another approach where policies are judged in terms of how they contribute to **human capabilities**. According to Sen (p.10, 1999), five distinct types of rights and opportunities feature in human capabilities: political freedom, economic facilities, social opportunities, transparency guarantees, and protective security. In practice, human capabilities are measured through metrics such as the United Nations' Human Development Index (in fact the creation of this index was inspired by Sen's research).

VI. Other Important Considerations for Policymaking

There are several other important factors that need to be considered while evaluating economic policies beyond selecting the appropriate economic model, the appropriate utility functions for the people in the model, and the appropriate measure of social wellbeing.

1. Whom to include in the measure of social wellbeing and how much weight should be accorded to different individuals.
 - a. Should a policy maker only care about the impact of policy on citizens of her own country or should she also consider the impact of the policy on citizens of other countries? What about immigrants, emigrants, refugees, etc.? ⁸
 - b. Should a policy maker care the same about all the people or should she prioritize some? For instance, should a policy maker care more about young people or more about old people? What about people who are important from a political perspective (e.g., people who live in swing districts in the United States)?
 - c. Should a policy maker care about future generations that are yet to be born? Should she prioritize these unborn generations or the currently living? This is related to questions about the appropriate time discount rate and is an important consideration for topics in environmental economics and other long-term problems – many of those who will suffer the worst consequences of climate change are not yet born.

⁸ Even when a policymaker in a particular country ignores the utility values of foreigners in her chosen social welfare function, she might need to *indirectly* care about citizens in a foreign country for two possible reasons: 1. The economic circumstances of foreigners in the other country might have a direct impact on the economic outcomes for the citizens in her home country (for instance, through trade opportunities). 2. The citizens of the home country might care about the wellbeing of the foreigners; maximizing social welfare of the home country citizens might therefore necessitate considerations of the wellbeing of foreigners too.

2. Other notions of justice and fairness could be important to consider too. For example, should a policy maker consider historical wrongs? The developed world is largely responsible for the accumulation of greenhouse gases in the atmosphere. Do they therefore bear a greater responsibility for making economic sacrifices to reduce the brunt of climate change?
3. On a related note, while economics frameworks are best suited for consequentialist analysis, policymakers might also have **deontological** considerations. Deontology is a normative ethical theory that holds that actions (or policies) need to be evaluated based on fundamental moral principles rather than in terms of their potential consequences. In the context of policymaking, some policymakers might be committed to fundamental human rights and liberties, international agreements like the Geneva conventions, or national legal principles enshrined in the constitution, etc., for fundamental reasons that extend beyond the material usefulness of those commitments. To discuss the nuance involved, a few examples might be helpful.
 - a. Among those discussing economics issues, especially online, there are several libertarians who believe that human liberty is of paramount importance. *Consequentialist* libertarians argue for their position based on the belief that good outcomes will result from the implementation of libertarian policies. *Deontological* libertarians, on the other hand, argue for their position in a more fundamental manner; they believe libertarianism is the only morally just way to organize an economy even if there were some other economic arrangements that would yield better economic outcomes.
 - b. Similarly, Senator John McCain, a staunch opponent of torture as an interrogation technique, often argued on moral grounds. According to John McCain, torture was fundamentally wrong and therefore unacceptable. In his view, therefore, any argument about the effectiveness or ineffectiveness of torture as an interrogation technique was irrelevant to the fundamental immorality of the practice.
 - c. John Rawls, after whom the Rawlsian welfare function is named, argued that first and foremost, “Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all” (Rawls, p. 220, 1971). Thus, according to Rawls, consequentialist considerations were to only be undertaken after a deontological bestowing of equal basic liberties. Rawls referred to this as the first principle.

References

- Arrow, K.J., 1963. Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review* 53, 941–973.
- Backhouse, Roger E., and Steven G. Medema. 2009. “Retrospectives: On the Definition of Economics.” *Journal of Economic Perspectives* 23 (1): 221–33.
- Box, G. E. P. 1979. “Robustness in the Strategy of Scientific Model Building.” In *Robustness in Statistics*, edited by Robert L. Launer and Graham N. Wilkinson, 201–36. Academic Press.

Rawls, J., 1999. A Theory of Justice: Revised Edition. Harvard University Press.

Sen, A., 1999. Development as Freedom, Revised Edition. Alfred Knopf.