

BOUNDED RATIONALITY AND ENDOGENOUS PREFERENCES

Robert Östling



EFI THE ECONOMIC RESEARCH INSTITUTE

Introduction

The standard approach in economics takes preferences as given and assumes that economic agents make rational decisions based on those preferences. In this thesis I relax both assumptions in a number of economic applications. These applications are somewhat diverse and belong to different fields of economics—game theory, microeconomics and political economics—but they can all be viewed against the backdrop of recent developments in behavioral economics.

The growth of behavioral economics took off in the late 1990s as a critique of neoclassical economics, which generally viewed man as self-interested and rational. Although few economists probably wholeheartedly ever believed that all of us are economic men, until recently, the core of economic theorizing relied mainly on that view of human behavior. This has changed and there are now plenty of economic models based on the alternative assumptions that individuals are imperfectly rational or care for others. The first generation of papers in behavioral economics focused on convincing other economists that self-interest and perfect rationality are not always the best descriptive assumptions. The second generation of papers in behavioral economics explores the implications of those alternative assumptions about human behavior in various economic settings. I hope that most of the papers in this thesis belong to the second generation and I think this is most clear in the first paper of this thesis.

In the first paper, which is written jointly with my much appreciated advisor Tore Ellingsen, we use a model of bounded rationality in order to better understand a classical question in game theory, namely how communication affects behavior in strategic interactions. Thomas Schelling discussed communication in games already in 1960 in his book *The Strategy of Conflict*. He described the cold war between the US and the Soviet Union as a simple two player game (see Chapter 9 in *The Strategy of Conflict*). Either side could choose to start a nuclear attack or not to attack. Both would be better off if no side attacked, but if one side attacks, the other side would like to do so as well. Game theorists recognize this game as a Stag Hunt—named after Jean-Jacques Rousseau’s description of a hunt—a game that has two equilibria: one good, but risky equilibrium where both sides stay calm, and a bad equilibrium in which both sides attack. The most interesting—and challenging—question with this game is to understand which equilibrium that will prevail and whether communication between players can help them to obtain the good outcome. In retrospect, we know that there was no nuclear war, which may partly be due to the existence of the hot line between Moscow and Washington. In the last chapter of *Arms and Influence* from 1966, Thomas

Schelling discussed the hot line more in depth and concluded that communication increases the chances that a nuclear war can be avoided: “The hot line is not a great idea, only a good one. (p. 262)”

Imagine yourself in the shoes of John F. Kennedy and that you receive a call from Nikita Khrushchev who tells you that today he and his close friend Fidel Castro is *not* going to start a nuclear war against you. Should you believe him? Initially, the consensus among game theorists was that you should since the message is *self-committing*, that is, if Mr. Khrushchev thinks that you believe the message, it is in his best interest to do what he said and not attack. However, Robert Aumann pointed out early on that Mr. Khrushchev’s message is not *self-signaling*. If he for some reason has decided to start a nuclear war, although that is a worse outcome than no war, then he would have told you that he wasn’t going to attack (since a surprise attack is better than a full scale nuclear war right from the start). Although this line of reasoning seems theoretically sound, experiments have shown that communication often is successful in getting people to coordinate on the good outcome.

One of the findings in the first paper is that when players are boundedly rational it makes sense for President Kennedy to believe in Mr. Khrushchev. The notion of bounded rationality used in the first paper is based on the steps of reasoning that many people do when they think about how to act in strategic situations. The kind of thinking that we model goes roughly along the following lines:

What if Khrushchev is completely irrational? Most likely, he’s going to naïvely say what he is going to do, so I might as well believe him. But what if he is not completely irrational? Then he would figure out that I would believe his message, so he’s going to be truthful and I better believe him.

This way of thinking about bounded rationality goes under the name of level- k reasoning and has been used to explain behavior in a wide range of experiments. The first paper applies that model to communication in games and argues that it provides a better account of how real human beings communicate than the perfectly rational model.

The second paper focuses on a different and much debated question in game theory, namely to what extent mixed equilibria are reasonable descriptions of behavior. In a mixed equilibrium, players attach probabilities to (some of) the strategies they have available and randomize based on those probabilities when they play. Typical games which have realistic mixed strategy equilibria are penalty kicks in soccer, Matching Pennies and Rock-Paper-Scissors. Although these are simple two player games, they share some features with the game studied in the second paper, which has thousands of players and strategies.

The second paper is written together with Colin Camerer, who generously hosted me during my stay at California Institute of Technology, as well as Joseph Tao-yi Wang and Eileen Chou. In the paper, we study the LUPI game that was introduced by the Swedish gambling monopoly in 2007. The rules of the game are simple: each person picks an integer between 1 and 99,999 and the person that picked the lowest unique number, in other words, the lowest number that was only picked by one person,

wins a fixed prize. The mixed equilibrium of this game is that each player plays 1 with highest probability and attaches a lower probability the higher the number is. The idea of playing lower numbers with higher probability is intuitive, but the exact magnitude of these probabilities is not (judge for yourself by looking at Figure 1 on page 52). Although the equilibrium is both difficult to compute and not particularly intuitive, players quickly learn to play close to the equilibrium prediction. To corroborate our findings, we also run classroom experiments in which students played the LUPI game with very similar results.

What more is there to say about the LUPI game if people play according to the equilibrium prediction *as if* they were perfectly rational? In my opinion, a theoretical model should not only be judged by its predictions, but also by the soundness of its assumptions. The assumptions that underlie the equilibrium prediction requires a great deal in terms of both rationality and computational power and we would therefore like to have a theory with more realistic assumptions that explain how people learn to play close to the equilibrium prediction. The answer we provide turns out to be remarkably simple: If people simply imitate numbers around previous winning numbers, they will soon learn to play something which is very similar to the equilibrium prediction. This learning dynamic requires almost no rationality of the players.

The final piece of the LUPI puzzle is to account for how people play the game the first time they play it, before they have had any opportunity to learn. Primarily to explain behavior in early rounds, we develop a model based on a similar notion of bounded rationality as in the first paper: the most naïve players pick numbers completely randomly, players that do one step of reasoning pick very low numbers and those that do two steps of reasoning therefore pick slightly higher numbers (continuing in a similar fashion for more steps of reasoning). This model combined with the learning model can account for how players play initially and then gradually learn to play close to the equilibrium prediction.

In the first two papers I try to develop more realistic descriptions of human behavior by relaxing the rationality assumption. In the third and fourth papers, I instead relax the assumption that people have stable and exogenous preferences. In some circumstances it is a valid simplification that preferences are exogenous, but in others it is not. Preferences do change and they sometimes do so in predictable ways, and that may have economic implications.

One area in which I believe preference changes to be of particular importance is with respect to moral preferences. The third paper focuses on moral preferences related to consumer goods. The paper builds on the psychological theory of cognitive dissonance which briefly stated says that whenever we experience contradicting cognitions, we experience a negative feeling that we are motivated to reduce. For example, if you are concerned about climate change, you might feel bad when you think about that you ought not to travel by plane at the same time as you buy a flight ticket. In order to reduce that feeling of dissonance you may rationalize the consumption decision, for example by convincing yourself that this particular trip is morally motivated.

In the paper, cognitive dissonance is combined with standard consumption theory. A consumer decides how to allocate his income between immoral and moral goods.

A lower price, or higher income, might be a temptation to buy more of the immoral good, which is assumed to be in conflict with the consumer's moral values. In order to reduce the resulting dissonance, the consumer rationalizes his consumption decision by changing moral values. For example, a decrease in the price of air travel may not only increase air travel, but may also lead to consumers adapting their moral values and becoming more tolerant towards going by air. The paper also contains an empirical analysis which shows that we are more tolerant toward goods and activities we tend to consume more. For example, rich people are more tolerant toward tax evasion than poor people, whereas poor people are more tolerant toward benefit fraud than rich people are.

The fourth paper in this thesis, written together with my colleague and friend Erik Lindqvist, concerns group identification. Like the third paper, it focuses on how people's motivation might change in response to changes in the environment they face. People tend to identify with groups, be it their ethnic group, gender, company or neighborhood, which in extreme cases can lead to violent conflicts. For economists, a natural way to understand group identification is that people join certain groups to form coalitions in order to extract more of some material resource. However, there are plenty of experiments, particularly by social psychologists, that suggest that the tendency to identify with groups is more fundamental and not always motivated by material incentives. The fourth paper incorporates some of the insights from social psychology into economic theory in order to better understand the determinants of the level of redistribution from rich to poor. We focus primarily on the interaction between ethnicity, social class and redistribution, which has interested social scientists throughout the 20th century. In some ways, we formalize ideas that go all the way back to Gunnar Myrdal's *An American Dilemma* and other scholars that have pointed out the black-white racial relationship as the reason for the difference between the US and Western Europe when it comes to redistribution.

In the simplest version of our theoretical model, individuals belong to one social class, rich or poor, as well as one ethnic group, black or white. Individuals choose whether to identify with their class or ethnic group, and that choice in turn determines their preferences and how they cast their vote over redistributive policies. For example, a poor white person in the model chooses between identifying with the white or with the poor. If he identifies himself as white, he becomes altruistic toward the white group which contains both rich and poor whites. If he instead identifies himself as poor, he becomes altruistic towards the poor. This means that he supports lower levels of redistribution from rich to poor if he identifies with the white group than if he identifies with the poor. There are two kinds of equilibria in the model. In the "European equilibrium", the poor identify as poor and favor high taxes and the level of redistribution is high. In the low-tax "US equilibrium", the poor whites identify with the white and redistribution is low.

An implication of the model is that an increase in the size of an ethnic minority, for example as a result of immigration, might lead to the ethnic majority switching to identifying with their ethnic group, which reduces the level of redistribution. This

is in line with several empirical studies that have found that more ethnically diverse societies have lower levels of redistribution.

The third and fourth papers both focus on how the social and economic environment affects people's preferences. Both papers imply that preferences are likely to be heterogeneous across individuals. For example, the rich are more tolerant toward tax evasion than the poor and someone who belongs to a poor ethnic group is likely to prefer more redistribution than an equally poor person that belongs to a rich ethnic group. In the fifth and final paper, which is also written jointly with Erik Lindqvist, we take preference heterogeneity as given and study its economic implications. More specifically, we empirically study the relationship between polarization of citizens' preferences and the size of government.

Why should we expect a relationship between polarization and government size? Suppose that you live in a heterogeneous society in which people have widely different ideas about what the most appropriate policies are. In such a society, it is quite likely that the policies the government implements will differ from your preferred ones. Irrespective of your own ideological position, you are therefore likely to prefer a smaller government the more polarized the society is.

To test these ideas, we derive a measure of the level of polarization in a country based on responses to survey questions about economic policy. We show that there is a strong negative relationship between political polarization and the size of government. The more polarized a country is, the smaller is the government. The effect is only present in the most democratic countries and the results are therefore consistent with a political mechanism like the one just described.

The remainder of this thesis consists of the five papers introduced above. The papers are self-contained and written with the purpose of eventually being published as separate articles in scientific journals. Although the topics covered are disparate I hope that this introduction has inspired you to continue reading the parts that interest you the most.



There is a long list of people without whom this thesis would have been in much worse shape. My co-author and advisor, Tore Ellingsen, has played a particularly important role. My other co-authors have also been crucial: Erik Lindqvist, Colin Camerer, Joseph Tao-yi Wang and Eileen Chou. Although Magnus Johannesson is not (yet) a co-author, he has provided much support throughout my graduate studies. There are many other colleagues, friends and family members that have played different and important roles for me and this thesis, but they are too many to be mentioned here and they deserve more attention than can be given on a few lines in this thesis. Stockholm School of Economics and the Jan Wallander and Tom Hedelius Foundation paid for the five inspiring, enjoyable and productive years.