

A Preference for Selfish Preferences

The Problem of Motivations in Rational Choice Political Science

Ignacio Sánchez-Cuenca
Juan March Institute, Madrid, Spain

This article analyzes the problem of preference imputation in rational choice political science. I argue against the well-established practice in political science of assuming selfish preferences for purely methodological reasons, regardless of its empirical plausibility (this I call a preference for selfish preferences). Real motivations are overlooked due to difficulties of imputing preferences to agents in a non-arbitrary way in the political realm. I compare the problem of preference imputation in economic and political markets, and I show the harmful consequences of the preference for selfish preferences in the field of collective action.

Keywords: *rational choice theory; preference; functionalism; collective action*

1. Introduction

Since its origins, political science has imported to the study of politics the methodology and approach of other social sciences. Political science, as such, has never developed its own distinctive approach. Rather, it has examined political institutions and political behavior borrowing analytical tools and methods from other disciplines such as law, anthropology, social psychology, sociology, and, more recently, economic theory.

The economic approach to politics faces some difficulties that stem from the inherent ambiguity about the goals that political actors pursue. It is not clear what assumptions should be made about the preferences of the actors. Given this ambiguity, it is really hard to dispel the impression that the

imputation of preferences to the agents is ad hoc. Most political scientists try to avoid the problem, employing a narrow view of self-interest: agents are moved by selfish, material motivations. This is what I call a preference for selfish preferences. This preference is not grounded on philosophical or empirical grounds, because both philosophers and empirical social scientists have shown the variety of human motivations. The preference for selfish preferences is rather based on methodological reasons. The assumption of narrow self-interest is accepted simply because it makes certain kind of explanations that are favored by the economic approach possible.

The result is awkward: highly sophisticated models based on implausible assumptions about motivations. The evidence on motivations that are not reducible to narrow self-interest is dismissed. One curious consequence of the attempt to explain every political action in narrow rational choice terms is the emergence of a new and strange type of functionalism: The agent chooses certain course of action because it increases her utility, even if the agent is not aware of that increase.

The aim of this article is to analyze the methodological preference for selfish preferences in political science of the rational choice persuasion. I do not claim that there is something intrinsically wrong in economic theory, or that the economic approach cannot be used beyond economic activity. My point is only that the economic approach in political science is distorted by this methodological preference. I do two things: First, I try to understand the reasons why the preference for selfish preferences is so widely shared; second, I provide some arguments about the necessity of going beyond the selfish assumption.

This article takes sides in the ongoing discussion about the role of rational choice theory in political science. The discussion has not been exclusively restricted to the domain of ideas. In 2000, a group of American scholars, acting under the collective name of "Perestroika," protested against the hegemony of formal and quantitative political science. They published a collective letter in the *New York Times* and obtained some concessions within the *American Political Science Association* (Monroe 2005). Yet, I am interested here only in the intellectual debate (Brennan and Hamlin 2000; Cox 1999; Friedman 1996; Goldthorpe 2000; Green and Shapiro 1994; Lewin 1991; Lovett 2006; Mackie 2003; MacDonald 2003; Satz and Ferejohn 1994; Shapiro 2005).

The question in dispute can be posed rather crudely: What have we learned about politics thanks to rational choice theory? According to critics, very little: Formal models are highly artificial, their creators do not care much about empirical testing, and when they do, they often commit methodological sins such as arbitrary domain restriction, adhocness, and lack of clarity about the null hypothesis. Critics claim that the only models

that have some empirical backing are those that are obvious and whose conclusions were known for long.

The analysis of collective choice in committees and elections is a perfect illustration of a body of theoretical literature that is far removed from empirics. Basically, the formal literature establishes that when decisions are made on a single dimension (e.g., the left-right or liberal-conservative dimension), the only equilibrium is the position of the median voter (Black 1958; Downs 1957). However, in multidimensional spaces, there is disequilibrium. The chaos theorems or the results on manipulation of the agenda predict arbitrariness and cycles (Riker 1982). According to this research program, if we seldom observe cycles or instability, it is because, assumedly, institutions fix them. This literature is particularly speculative. To begin with, models do not specify what problems are dealt with in multidimensional versus single dimensional spaces. Most research, both on Parliaments and electorales, shows that political issues are usually reduced to a single underlying dimension. Besides, the theory is based on entities (cycles, manipulation) that are hard, if not impossible, to observe. The evidence used to persuade about their existence is mostly anecdotal, and, as Mackie (2003) has shown convincingly, these anecdotes are empirically dubious. Lastly, the idea that institutions exist to solve the problems created by disequilibrium has some functionalist flavor: Those who create the institutions are not necessarily aware of the reasons why they do so (more on rational choice functionalism below).

Generally speaking, a rational choice theorist could answer acknowledging that all these criticisms were fair, but only with regard to the first waves of rational choice. Today, goes on the reply, more sophisticated models and greater concern about empirics have overcome many of these criticisms. In fact, in the last 10 years or so, we have witnessed an impressive integration of formal modelling and empirical testing. There are some landmarks worth mentioning: Boix (2003), Epstein and O'Halloran (1999), Huber and Shipan (2002), or Tsebelis (2002) are paradigmatic examples of formal models combined with rigorous econometric testing in comparative politics. Moreover, there have been serious attempts to reinforce the empirical side of formal models, like the EITM program ("Empirical Implications of Theoretical Models," see Aldrich, Alt, and Lupia 2008).

Yet, I am going to argue that solving the problem of empirical irrelevance is not only a matter of improving the formation of political scientists, so that they learn to combine models and empirics in a more fruitful way, paying more attention to the testable consequences of models. There are deeper causes for the lack of realism of many models in political science. In my view, the heart of the matter lies in the assumptions about motivations.

Formal models cannot work in politics unless the discipline makes an extra effort to consider preferences that are not merely selfish, material ones.

Curiously enough, economic theory has made serious progress in the last 15 years or so by incorporating non-standard motivations in its models (see Sobel 2005 for an excellent review). It is as if political science were more catholic than economic theory itself. Due to philosophical or methodological obfuscation, rational choice political scientists have paid too little attention to models that take into account moral principles, other-regarding preferences, or norms of reciprocity and fairness. If the rational choice approach in political science has not made progress in this direction, it is not because it is unfeasible, as shown by behavioral economics (Camerer, Loewenstein, and Rabin 2004), but because of methodological caveats regarding the imputation of preferences.

The article is divided into three sections. In section 2, I discuss the reasons why it is more difficult to impute preferences to political agents than to economic ones. Whereas the institutional structure of the market induces the preferences of entrepreneurs and consumers in a natural way, the institutional structure of the political game is too thin to induce preferences. To illustrate the argument, I comment on theories of party competition. Then, in section 3, I examine the philosophical and methodological reasons that have been provided to defend the necessity of a narrow view of self-interest, and how this has produced a sort of “rationalist functionalism.” Finally, in section 4, I show how a particular field of study, that of collective action, has been specially hindered by bad solutions to the problem of preference imputation.

2. Preferences In and Out of the Market

In its most abstract formulation, rational choice theory is agnostic about the motivations of people. Even if, historically, preferences were taken to be the bearers of interests, desires, or wants (Lewin 1996), the fact is that the formal apparatus of the theory has deprived the concept of any substantive content. A preference is just a binary relation over alternatives that satisfies certain consistency properties (Sen 1970, Ch.1). The theory, at this level of generality, does not have much to say about the reasons why alternative x might be preferred to alternative y : This could be the case because the agent has a stronger desire for x than for y , because other people prefer x to y , or because x accords with some duty, whereas y doesn't. Here, the concept of preference is devoid of any substantial meaning, to the point that it can cover even choices made under the force of binding obligations.

When the theory tries to find empirical applications, it does not employ such an empty notion of preference. Typically, preferences are taken to represent the self-interest of the agent, self-interest being interpreted in a rather narrow and materialistic way. Otherwise, rational choice theory becomes merely a logic of choice without empirical content. Few observational consequences can be drawn from the theory if the reasons for preferring one alternative over the other are left completely unspecified.

The heart of the matter lies in the private nature of preferences. Supposedly action is to be explained in terms of the preferences of the agent. But preferences are not observable. How can they have, then, explanatory power? Everything hinges upon preferences, but we have to impute preferences to agents because we do not have direct access to them. If the imputation is not severely constrained, the theory becomes unfalsifiable: It is just a matter of finding a preference order that fits the pattern of observed actions. If there were no restrictions on the preferences that can be imputed to the agent, the theory would be true by construction. The consistency restrictions of the pure theory of choice (see, for instance, Sen 1970) are too weak to solve the imputation problem.

Some solutions in economics have been indeed drastic. For instance, Becker (1976, Ch.8) proved that the demand law can be obtained without rationality assumptions. And Becker and Stigler (1977) proposed to keep preferences fixed, explaining behavior in terms of differences of capital. But, as Elster (1997) has convincingly argued, very similar problems arise regarding some forms of unobservable capital.

A more modest proposal is to focus on the institutional context in which the action occurs. The context is sometimes sufficient to determine certain course of action regardless of the ultimate preferences of the agents. The market is a good case in point. We can explain the behavior of the entrepreneur referring to the aim of making profit. But, as Ferejohn and Satz (1994, p. 79) note, profit-maximizing behavior is compatible with "an other-worldly interest in eternal salvation or a crude materialism, or a propensity for innovation and risk taking." What entrepreneurs want profit for does not have to enter into the theory about the entrepreneur's market behavior.

Now, why does the entrepreneur have to maximize profits? Two typical answers run as follows. First, an evolutionary one: Only those entrepreneurs who maximize profits survive in a competitive market (Alchian 1950). Second, an institutional one: Being entrepreneur is a role institutionally determined that consists of running risks to gain benefits. If the economic activity of some agent does not fit this description, then we simply refuse to refer to this agent as an entrepreneur. The preferences of the entrepreneur are ultimately determined by its social role within the market.

Although the case about the consumer is slightly more complex, the institutional structure of the market also makes the ultimate preferences of consumers irrelevant. As Margolis (1982, 11) says, “if Saint Smith gives bread to the poor, then the benefited poor do not enter the market, but Smith’s preference map incorporates the demand for bread to give to the poor. We have no more need to distinguish between the bread Smith buys to give to the poor and that which he buys for his own consumption.” Demand for bread, the price system, and the budget constraint are sufficient to determine Smith’s choice regardless of the purpose for which Smith wants the bread.

Out of the market, it is an open question whether the institutional context will be thick enough to make ultimate preferences irrelevant. The comparison between the economic market and the “political market” in a democratic system is telling. The rational choice theory of political competition tries to understand elections as a market in which there are two types of agents, politicians, and voters, who are the counterpart to entrepreneurs and consumers respectively.

Let us start with politicians. Just as entrepreneurs seek benefits, politicians seek power. This was, of course, Anthony Downs’ (1957) assumption in his path-breaking economic approach to democracy. **Unlike the case of the entrepreneur, the reasons why politicians seek power are important to explain their behavior.** It matters whether politicians seek power to implement certain policies or to enjoy the rents associated with power (money, prestige, influence, or fame). It is true that in Downs’ model, the ultimate preferences of politicians (what do they want power for?) are inconsequential because of the restrictive assumptions he makes: Competition takes place in a single dimension under perfect information. However, when politicians are uncertain about the location of the median voter, divergent platforms are an equilibrium if politicians care about what policies are made, whereas there is convergence (the same policy positions) when politicians only value the rents of office (Alesina 1988; Calvert 1985; Wittman 1973; 1983). Office- and policy-seeking politicians are not the only possible types. In some models, politicians are ideologically dogmatic: They value more the ideological line of the party than power (Roemer 2001; Sánchez-Cuenca 2004). If parties were dominated by this kind of dogmatic politicians, there could be divergence even with full certainty about the median voter.

Regarding voters, the analogy with consumers is again problematical. On the one hand, the counterpart of prices in the political market is unclear. The cost of voting is constant, independent of what party is voted for by the voter. On the other hand, using the terminology of Brennan and Lomasky (1993, 15), the consumer is decisive, because the chooser obtains what she

chooses. By contrast, the voter is non-decisive. The voter does not choose between parties: The voter votes for some party, but the final outcome is not determined by the choice of this single voter (this gives rise to the paradox of voting, see Blais [2000] for a review of a gigantic literature). Given that the voter is not responsible for the final electoral outcome (her vote is one among many, usually thousands or even millions), many considerations may enter in the voter's decision, from ethical principles to purely selfish interests. This creates some indeterminacy in the motivations that can be imputed to voters. Thus, the lack of a well-articulated theory about voting behavior should be no surprise. **The motivations of voters are so varied that no single theory can aspire to cover all possible cases.** There is not room for a theory comparable in generality and parsimony to consumer theory.

As this brief review shows, different motivations about politicians and voters lead to different predictions about party competition. We are thus forced to make non-obvious assumptions about the preferences of these actors. How can we avoid arbitrariness in the imputation of preferences?

3. Narrow Preferences and Rational Choice Functionalism

Given that the institutional context rarely solves the preference problem out of the market, rational choice theory in political science tries to solve the problem assuming that political actors pursue narrow self-interest, understood as selfish, material interests. This is what I have called a preference for selfish preferences.

Assuming selfish preferences is justified in methodological terms. Due to the indeterminacy of preferences out of the market, any assumption that makes it easy to explain the facts is automatically deemed a suspicious one. Non-standard preferences are likely to produce, according to this argument, ad hoc and tautological explanations. Let us take the case of turnout. We can understand the decision to vote as an investment or as a consumption act (Ferejohn and Fiorina 1974). Most authors assume that the turnout decision is an act of investment (an instrumental action) simply because the alternative seems obvious and vacuous. The alternative is to consider that voting is done for its own sake: because it is fun in itself or because people think that it is their duty to vote. Some authors consider that voting is not that different from applauding in a concert or cheering the home team (Aldrich 1997, 387). However, if voting is explained in terms of a taste for voting, it seems that we have not made much progress in theoretical terms.

This methodological justification is indeed salient in the founding texts of rational choice political science. In *The Logic of Collective Action*, Mancur Olson contemplates in a long footnote the possibility of moral motivations that could lead the agent to contribute to a collective good to avoid the sense of guilt or the loss of self-esteem that would be produced by no cooperation when others cooperate. But he rejects this possibility for three different reasons, of which the first two are of particular relevance for the present discussion:

First, it is not possible to get empirical proof of the motivation behind any person's action; it is not possible definitely to say whether a given individual acted for moral reasons or for other reasons in some particular case. A reliance on moral explanations could thus make the theory untestable. Second, no such explanation is needed, since there will be sufficient explanations on other grounds for all the group action that will be considered. (Olson 1965, 61)

Given that preferences are not observable, the smaller the role of preferences in the explanation, the better. Or, to put it in another way, **moral motivations are superfluous as long as a general explanation in terms of selfish motivations is available.** A theory in which everyone has selfish preferences is superior to a theory in which some people are assumed to have moral motivations and some others non-moral ones. Olson's reference to "sufficient explanations" is interesting: He seems to mean explanations that fit the evidence even if they do not reflect faithfully the real motivations of agents.

When Downs discusses the motivations of voters, he makes a similar observation:

If it is rational to vote for prestige, why is it not rational to vote so as to please one's employer or one's sweetheart? Soon all behavior whatsoever becomes rational because every act is a means to some end the actor values. To avoid this sterile conclusion, we have regarded only actions leading to strictly political or economic ends as rational (Downs 1957, 263).

Again, Downs does not show that the motivations he rejects do not play a role in the choice of voters. **He considers rather that the methodological problems created by these motivations justify a decision to ignore them.**

Finally, William Riker argues in the same terms in his theory of coalitions:

So if the rationality condition is to be useful in models of behavior, one is forced to go back to the cruder and already somewhat discredited notion of an economic or political man who maximizes a utility that is scaled about the same way as money or power (Riker 1962, 20).

In these three quotes, motivations are arbitrarily narrowed for the sake of methodological requirements; otherwise, the argument runs, the theory loses its predictive capacity and becomes true by mere fiat. The ontological question, however, is not even considered. Do individuals really act on reasons that cannot be reduced to selfish self-interest?

This lack of attention to the basic ontological question about the motivations of the people has led to the emergence of a strange form of functionalist explanation in which the entity that benefits from certain kind of behavior is not a group (the species, society, a social class), but the person. Elster (1985, 27) argues that whereas in intentional explanation, the *explanans* is the *intended* effects of the action, in functional explanation, it is the *actual* effects, regardless of whether the agent had the intention to bring about these effects or not. Of course, if the agent is not aware of the actual effects, then we need a mechanism that links the action with its actual effects (natural selection, reinforcement; see Van Parijs 1981).¹

In rational choice functionalism, the actual consequences may play the role of hidden motivations of which the agent is unaware. If the agent carries out an action whose actual consequences are an increase of utility, but the agent did not choose this action because of these actual consequences, and the explanation is based on these consequences, we seem forced to conclude that the agent is moved by motivations that are not transparent to her. How these hidden motivations have explanatory power is an intriguing question that functionalists do not care to answer.

In Merton's classical terms, rational choice functionalism operates displacing manifest functions into latent ones. If the agent is unaware of the actual effects of her action, then these effects, if they are to have any explanatory power, cannot be framed into the purposive language of standard rational choice. If utility maximization is a latent function, then we need a non-rationalist mechanism that, obviously, cannot be provided by standard rational choice theory.

Perhaps this point can be made clearer with an example. Let us take Downs' (1957) hypothesis about the role of political ideology. In his view, ideologies exist because they save information costs to voters. If voters had to read carefully electoral platforms to determine which party best satisfies their interests, voters would have to spend many hours obtaining information about the policies that each party proposes and the likely consequences

1. About the contentious possibility of functional explanations without mechanisms, see Cohen (1978, Ch.9).

of these policies. Given that a single vote counts little in the determination of the final outcome, people do not have incentives to read all these materials. Yet, Downs argues, voters rely on ideology as a shortcut to anticipate what parties are going to do in case they get into power. Ideology, therefore, is helpful or beneficial for voters. It saves information costs.

This may seem an entirely logical explanation. Nevertheless, Downs does not make much effort to show that information costs are empirically related to the adoption of ideological beliefs. That is only a hypothesis that is never tested. Those that embrace Downs' theory are happy to proceed in "as if" terms: When voters make decisions in ideological terms, they act "as if" they were minimizing information costs. Given the absence of any evolutionary or reinforcement mechanism that accounts for the "as if" rider, the resulting functionalist explanation is untenable.

In fact, it makes sense to invert the structure of the explanation: It might well be that people get information about politics because of their ideological beliefs, and not the other way around. People with ideology develop an interest in politics. Ideology, so to speak, generates the incentives to obtain information about politics. Just as people with passion for soccer find it rewarding to learn about soccer players or about past matches, people with ideological beliefs enjoy learning about parties and candidates. Here, the acquisition of information is a consumption act, not a investment one (Fiorina 1990, 337). It is perhaps true that this alternative explanation is not particularly enlightening, but at least it is not based on highly questionable assumptions.

Because of the methodological reasons aforementioned, rational choice theory privileges explanations based on strict assumptions about selfish, materialistic preferences. The greater challenge lies in reducing apparent altruist or moral behavior to selfish motivations. Many rational choice theorists consider that the less obvious the explanation of the action, the higher its explanatory value. This approach is driven by methodological considerations alone. Thus, any signal of good will or of commitment to a cause by the agent is to be explained by an array of analytical devices compatible with standard preferences. If someone risks her life or her career for a collective cause, the analyst will try to reduce her actions to motivations that are totally unrelated to the collective cause.

Think for instance of middle-class white students of elite universities that participated in the 1964 Freedom Summer helping blacks to register in the South (McAdam 1988). Several of them were killed, and many still remember the shock and grief they felt when they discovered through direct experience the living conditions of the blacks and the treatment they received from police and white dwellers. A typical rational choice explanation will

be suspicious about the relevance of commitment to the cause as an explanation of the behavior shown by those who participated in this campaign. It will look elsewhere: in potential side payments (flirting with other youngsters during the summer), in social pressure (the person participated because many of her fellows did too), or in political ambitions (participation in the campaign was only a step in a political career). Any of these stories is usually deemed more likely than an explanation based on the person's sincere commitment to the cause.

I am not saying that commitment to the cause (see Sen 2002, Ch.5) is sufficient to explain behavior. To begin with, commitment by itself cannot account for variance; that is, why among those with political convictions, some participated but some did not. Yet, it is also true that without serious commitment, the other payments (in form of social acceptance, side payments, or reputation) are not sufficient to induce participation in such risky action, for there usually are more secure ways to obtain these payoffs. Probably, the right explanation would have to combine commitment with other factors, such as being in the right social network or sharing experiences with some other people.

Game theory has been used and abused in the "deconstruction" of any apparent instance of non selfish behavior. The trick consists, briefly put, in showing that in the long run, the agent is better off behaving in a cooperative way even if she has standard selfish preferences. For example, in the Centipede game, people cooperate despite the fact that backwards induction predicts defection in the first round. There is ample evidence in the laboratory about that (McKelvey and Palfrey 1992). This is a powerful reason to consider non-standard preferences. However, the game theorist prefers to introduce incomplete information: As long as there is some uncertainty about the payoffs of the players, each player, although fully selfish, may try to imitate the behavior of someone with non-selfish preferences. The result is that players end up better off even by selfish standards (Kreps 1990, Ch.13). The key concept here is reputation. No matter how selfish or mean the agent is, it is in her long-term interest to develop the reputation of being a non-selfish person. Thanks to this reputation, the agent will be able to obtain the gains of trade that she could not have gained had it been obvious to everyone else that she was a selfish person.

The underlying philosophy is expressed very clearly by Calvert (2002, 591): "political behaviour that appears expressive or altruistic when viewed from a purely political perspective may be purely instrumental when seen in its wider social perspective." That is the research program of rational choice political science when applied to non-economic issues. It only

makes sense under the (methodological) assumption that any explanation based on selfish preferences is superior to explanations in which selfishness is relaxed.

In more philosophical terms, what I am questioning is this primacy of methodology over ontology. If people really act out of principles, moral obligations, convictions, or sincere concern for others, the methodological virtues of a theory that does not take into account these factors are of no avail. That theory could only aspire to provide very incomplete explanations. Common sense, folk psychology, and even simple introspection show that the range of motivations that move the agent to act is wider than standard self-interest.

In the next section, I show the perils of this reductionist strategy in a particular field of research, that of collective action. We will see how rational choice theories of collective action have refrained from considering non-standard preferences due to methodological prejudice.

4. Motivation and Collective Action

Theories of collective action have tried to avoid any sort of revision about the motivations for cooperation. In a collective action problem, what is good for all is not necessarily good for each. The tension between the collective and the individual is a matter of degree. The individual may feel tempted to consume the collective good without contributing to its provision. This is the famous free-rider problem. If others contribute to the collective good, the individual is better off when she does not contribute: All members benefit from the provision of the collective good, but only those who contribute pay the cost.²

To solve the collective action problem, Olson considered two factors, group size and selective incentives. In large groups, the only solution, according to Olson, lies in selective incentives. These refer to positive or negative payments that can be directed towards either those who cooperate or those who do not. Think of pension benefits or job promotion chances to those who become part of a trade union. These are clearly positive selective incentives. Or think of social ostracism as a negative selective incentive applied to those in a group that do not cooperate with others in a collective endeavor.

2. More technically, collective goods show non-rivalry in consumption (consumption by an agent does not preclude consumption by others). Thus, the marginal cost of consumption is zero. However, the production of the collective good is costly. Thus, there will be an under-provision of the collective good.

As it is well known, the focus on selective incentives produces some paradoxical results, for these incentives do not have to be connected with the collective goal that is pursued by the members of the group. Thus, a professional association may attract members by offering discounts in holiday plans even if these are totally unrelated to the goals defended by the association. While this is perfectly sound in interest groups that seek economic gains, it gives rise to absurd results when applied to political collective goods. Green and Shapiro (1994, 79) refer to a Christian fundamentalist that sees two demonstrations, one pro-abortion, the other anti-abortion, and joins the pro-abortion one because it offers greater material selective incentives (refreshments, sandwiches, or whatever). Since her individual participation is not going to change policy, the only factor relevant in her decision is the benefits that she may obtain by choosing one or another demonstration. If the argument seems counterintuitive, it is because in the political context cooperation only makes sense when there is some consistency between the collective goals pursued by the group and the ideas or values that the individual holds (Chong 1991, 33).

The thesis that economic selective incentives are not enough to induce cooperation relies on the assumption that there is some basic political motivation that goes beyond material gains. Wilson (1974, Ch.3) tried to capture this point by suggesting the existence of purposive selective incentives (the benefits the agent obtains by the mere fact of striving for a collective purpose) when the collective good is political in nature. However, the bulk of the literature has not explored this possibility, focusing rather on non-motivational solutions, such as the production function of the collective good (Granovetter 1978; Heckathorn 1996; Lohmann 1994; Macy 1991; Marwell and Oliver 1993; Schelling 1978).

The production function of a collective good models how individual contributions are aggregated.³ The form of the function is determined by the nature of the collective good. Depending on the features of the collective good, individual cooperation is more or less likely. The basic games that reflect the different possibilities of collective action (the Prisoner's

3. Perhaps the most popular production function is Heckathorn's (1996). This is his specification:

$$P(R) = 1 - (D/N)^F$$

D stands for the number of people who defect (don't cooperate), *N* is the size of the group, and *F* is the exponent that determines the function shape. The crucial parameter here is *F*: if *F* = 1, the production function is linear; if *F* > 1, the function decelerates (further individual contributions have negative marginal returns); if *F* < 1, the function accelerates (further individual contributions have positive marginal returns).

Dilemma, the Assurance game, the Chicken game, the Privileged game) stem from different types of collective goods.

The analysis of the production function seems to offer a promising line of research in which motivations play no significant role. Everything turns on the technical features of the collective good. However, this literature has not made much progress specifying empirically which collective goods are more conducive to cooperative behavior. We know that collective action may adopt different shapes, but it is hard to translate these formal results into testable hypotheses. More importantly, it is perfectly possible to derive the same formal results about different games being played changing not the features of the collective good, but the motivations of the agents (Sánchez-Cuenca 2007). Even if the collective good is such that it generates a Prisoner's Dilemma, variations in the intensity of non-standard utility (utility that incorporates considerations about fairness or about principles) may transform the underlying game into a game more conducive to cooperation.

In theories about the production function of the collective good, it is often the case that formal models only work on the assumption that there is an initial group of unconditional cooperators. They form the "critical mass" that triggers further cooperation by other agents, who are only conditional cooperators (people who are willing to cooperate as long as a sufficient number of others cooperate too). The existence of this critical mass is taken for granted. Yet, the existence of unconditional cooperators is problematical for any theory based on standard self-interest. Why should there be some people willing to cooperate regardless of what others do? In a small group, as Olson predicted long ago, this may make sense, but not in a large group. Motivations seem to enter here through the back door. Note the curious role of the critical mass: Independent of the features of the collective good, there is motivational variation among the agents. Some agents are more willing to cooperate than others. This, necessarily, implies that at least some agents have preferences that go beyond narrow self-interest.

The literature on collective action is badly divided between rational choice theories and empirical studies of social movements. The former are extremely rigorous from a formal point of view, but fail to address key issues about the role of motivation in the dynamics of collective action. The latter are more comprehensive, but they are not theoretically informed, even if they have created a framework and a vocabulary to describe how collective action is organized and developed (McAdam 1982; Tarrow 1993). Rational choice models of collective action cannot have empirical relevance unless they expand the range of motivations that can be imputed to agents and integrate some of the findings of the empirical literature.

5. Conclusions

The introduction of the economic approach in political science has transformed the discipline very deeply. Greater standards of rigor, higher theoretical sophistication, and more emphasis on explanation (as opposed to empirically rich description), are obvious consequences of this transformation. Yet, the reluctance of many political scientists to rational choice is not entirely unjustified. Very often, formal models are artificial and far removed from empirics. In my view, this is not due exclusively to the division of labor between basic and applied research. The irrelevance of much formal theory has deeper roots.

According to the argument I have defended in this article, the real problem lies in the problem of preference imputation. While in market transactions between producers and consumers, the institutional structure is such that the ultimate preferences of economic agents are irrelevant, in the political realm the issue of preference imputation looms large. The preferences of politicians or voters are unclear. This ambiguity creates some methodological troubles, as the critic may believe that the model works simply because the modeller has too much latitude to adjust the preferences to fit the observed behavior. To remove any suspicion about adhocness, or to avoid the accusation of tautology, formal modellers proceed to assume that political agents act out of selfish, materialistic motivations.

I have argued that, by doing so, rational choice political science privileges methodological concerns over ontological ones. The assumption of selfish, material preferences is too restrictive. Obviously, powerful theories always overlook many details of reality that are considered accidental or superfluous. But the agent's actual motivations cannot be relegated to that category. Motivations are what lead agents to act in a certain direction. They are the cornerstone of any intentional explanation.

Very few rational choice explanations in political science have gone beyond narrow self-interest (an interesting exception is Levi 1997). Given the methodological difficulties of incorporating non-standard preferences in models with applications in the realm of politics, the most natural reaction to the convoluted and artificial nature of rationalist models is to go for models of imperfect rationality (bounded rationality, evolutionary game theory, agent-based models). However, it seems to me that the importance of extreme behavior in politics (from suicide missions to joining revolution to volunteering for altruist causes) requires a further look at the possibility of enriching the range of motivations that is considered plausible in rational choice theory.

References

- Alchian, A. A. 1950. Uncertainty, evolution and economic theory. *The Journal of Political Economy* 58 (3): 211-21.
- Aldrich, J. H. 1997. "When is it rational to vote?" In *Perspectives on public choice*, edited by D. C. Mueller. Cambridge: Cambridge University Press.
- Aldrich, J. H., J. Alt, and A. Lupia. 2008. "Methodology and micro foundation." In *The Oxford Handbook of Political Methodology*, edited by J. M. Box-Steffensmeier, H. E. Brady, and D. Collier. Oxford: Oxford University Press.
- Alesina, A. 1988. Credibility and policy convergence in a two-party system with rational voters. *American Economic Review* 78 (4): 796-805.
- Becker, G. S. 1976. *The economic approach to human behavior*. Chicago: The University of Chicago Press.
- Becker, G. S., and G. J. Stigler. 1977. De Gustibus Non Est Disputandum. *American Economic Review* 67 (2): 76-90.
- Black, D. 1958. *The theory of committees and elections*. Cambridge: Cambridge University Press.
- Blais, A. 2000. *To vote or not to vote?* Pittsburgh: University of Pittsburgh Press.
- Boix, C. 2003. *Democracy and redistribution*. Cambridge: Cambridge University Press.
- Brennan, G., and A. Hamlin. 2002. *Democratic devices and desires*. Cambridge: Cambridge University Press.
- Brennan, G., and L. Lomasky. 1993. *Democracy and decision: The pure theory of electoral preference*. Cambridge: Cambridge University Press.
- Calvert, R. 1985. Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science* 29 (1): 69-95.
- . 2002. "Identity, expression, and rational-choice theory." In *Political Science: State of the Discipline*, edited by I. Katznelson and H. V. Milner. New York: Norton.
- Camerer, C. F., G. Loewenstein, and M. Rabin, eds. 2004. *Advances in behavioral economics*. Princeton: Princeton University Press.
- Chong, D. 1991. *Collective action and the civil rights movement*. Chicago: The University of Chicago Press.
- Cohen, G. A. 1978. *Karl Marx's theory of history: A defence*. Princeton: Princeton University Press.
- Cox, G. W. 1999. The empirical content of rational choice theory. *Journal of Theoretical Politics* 11 (2): 147-69.
- Downs, A. 1957. *An economic theory of democracy*. New York: Harper and Row.
- Elster, J. 1985. *Making sense of Marx*. Cambridge: Cambridge University Press.
- . 1997. Review of Gary S. Becker, *Accounting for tastes*. *University of Chicago Law Review* 64 (2): 749-64.
- Epstein, D., and S. O'Halloran. 1999. *Delegating powers*. Cambridge: Cambridge University Press.
- Ferejohn, J. A., and M. P. Fiorina. 1974. The paradox of not voting: A decision theoretic analysis. *American Political Science Review* 68 (2): 525-36.
- Fiorina, M. P. 1990. "Information and rationality in elections." In *Information and democratic processes*, edited by J. A. Ferejohn and J. H. Kuklinsky. Urbana: University of Illinois Press.
- Friedman, J., ed. 1996. *The rational choice controversy*. New Haven: Yale University Press.
- Goldthorpe, J. H. 2000. *On sociology*. Oxford: Oxford University Press.
- Granovetter, M. 1978. Threshold Models of Collective Behavior. *American Journal of Sociology* 83 (6): 1420-43.

- Green, D., and I. Shapiro. 1994. *Pathologies of rational choice*. New Haven: Yale University Press.
- Heckathorn, D. D. 1996. The dynamics and dilemmas of collective action. *American Economic Review* 61 (2): 250-77.
- Huber, J. D., and C. R. Shipan. 2002. *Deliberate discretion?* Cambridge: Cambridge University Press.
- Kreps, D. M. 1990. *A course in microeconomic theory*. Princeton: Princeton University Press.
- Levi, M. 1997. *Consent, dissent, and patriotism*. Cambridge: Cambridge University Press.
- Lewin, L. 1991. *Self-interest and public interest in western politics*. Oxford: Oxford University Press.
- Lewin, S. B. 1996. Economics and psychology: Lessons for our own day from the early twentieth century. *Journal of Economic Literature* 34 (3): 1293-1323.
- Lohmann, S. 1994. The dynamics of informational cascades. *World Politics* 47 (1): 42-101.
- Lovett, F. 2006. Rational choice theory and explanation. *Rationality and Society* 18 (2): 237-72.
- MacDonald, P. K. 2003. Useful fiction or miracle maker: The competing epistemological foundations of rational choice theory. *American Political Science Review* 97 (4): 551-65.
- Mackie, G. 2003. *Democracy defended*. Cambridge: Cambridge University Press.
- Macy, M. 1991. Chains of cooperation: Threshold effects in collective action. *American Sociological Review* 56 (6): 730-47.
- Margolis, H. 1982. *Selfishness, altruism and rationality*. Chicago: The University of Chicago Press.
- Marwell, G., and P. Oliver. 1993. *The critical mass in collective action: A micro-social theory*. Cambridge: Cambridge University Press.
- McAdam, D. 1982. *Political process and the development of black insurgency, 1930-1970*. Chicago: University of Chicago Press.
- McAdam, D. 1988. *Freedom summer*. Oxford: Oxford University Press.
- McKelvey, R. D., and T. R. Palfrey. 1992. An experimental study of the Centipede game. *Econometrica* 60 (4): 803-36.
- Monroe, K. R., ed. 2005. *Perestroika! The raucous rebellion in political science*. New Haven: Yale University Press.
- Olson, M. 1965. *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Riker, W. H. 1963. *The theory of political coalitions*. New Haven: Yale University Press.
- . 1982. *Liberalism against populism*. Prospect Heights: Waveland.
- Roemer, J. 2001. *Political competition: Theory and applications*. Cambridge, MA: Harvard University Press.
- Sánchez-Cuenca, I. 2004. Party moderation and politicians' ideological rigidity. *Party Politics* 10 (3): 325-42.
- . 2007. Cooperar por Principio. *Revista Internacional de Sociología* 65 (46): 11-35.
- Satz, D., and J. Ferejohn. 1994. Rational choice and social theory. *The Journal of Philosophy* 91 (2): 71-87.
- Schelling, T. C. 1978. *Micromotives and macrobehavior*. New York: Norton.
- Sen, A. 1970. *Collective choice and social welfare*. San Francisco: Holden-Day.
- . 2002. *Rationality and freedom*. Cambridge, MA: Harvard University Press.
- Shapiro, I. 2005. *The flight from reality in the human sciences*. Princeton: Princeton University Press.
- Sobel, J. 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* 43 (2): 392-436.
- Tarrow, S. 1993. *Power in movement: Social movements, collective action and politics*. Cambridge: Cambridge University Press.

- Tsebelis, G. 2002. *Veto players: How political institutions work*. Princeton: Princeton University Press.
- Van Parijs, P. 1981. *Evolutionary explanation in the social sciences*. Totowa: Rowman and Littlefield.
- Wilson, J. P. 1974. *Political organizations*. Princeton: Princeton University Press.
- Wittman, D. 1973. Parties as utility maximizers. *American Political Science Review* 67 (2): 490-98.
- Wittman, D. 1983. Candidate motivation: A synthesis of alternative theories. *American Political Science Review* 77 (1): 142-57.

Ignacio Sánchez-Cuenca is professor of political science (Juan March Institute, Madrid) and Associate Professor of Sociology (Universidad Complutense, Madrid). He has been Rice Visiting Professor at Yale University. His research interests are democratic theory, political violence, and methodology. He is co-editor with José M. Maravall of *Controlling Governments* (Cambridge University Press, 2008) and has published articles in journals such as *European Journal of Sociology*, *European Union Politics*, *Party Politics*, *Government and Opposition*, and others.