

The Weak Rationality Principle in Economics

GEBHARD KIRCHGÄSSNER^a

Keywords: Rationality, Self-Interest, Micro-Foundation, Bounded Rationality

JEL Classification: B41

1. Introduction

It is hardly disputed that the ‘rational choice’ or ‘rational actor’ approach, i.e. explaining human behaviour as *rational* behaviour under constraints, is more prominent in Economics than in all the other social sciences. This approach, of course, is also applied to other social sciences, like Political Science or Sociology, but in those it is far less dominant and much more disputed. There is much less unanimity about what ‘rationality’ means: the interpretations range from the strict application of expected utility theory as developed by VON NEUMANN and MORGENSTERN (1944) via full rationality and bounded rationality in the sense of SIMON (1955), the assumption that rationality simply implies “that agents are not in fact stupid” (LAGUEUX, 2004, p. 31), up to the weak rationality principle discussed in this paper. Along this line, the empirical content of the rationality assumption decreases: it is highest for the expected utility theory but zero, or at least quite close to zero, for the weak rationality principle.

The rationality principle is basically a descriptive one: its aim is to explain how people behave or act, respectively. Sometimes, however, it is also seen as a prescriptive one: it demonstrates what rational behaviour should look like; it provides a reference for rational behaviour. This holds in particular for the von Neumann–Morgenstern variant of expected utility maximising as applied, for example, in the Capital Asset Pricing Model of Finance Theory. Such a normative use is, however, less possible the weaker the rationality assumption is. The weak

a University of St. Gallen, Swiss Institute for International Economics and Applied Economic Research (SIAW-HSG), Leopoldina and CESifo. Mailing Address: University of St. Gallen, SIAW-HSG, Bodanstrasse 8, CH-9000 St. Gallen, Switzerland, Gebhard.Kirchgaessner@unisg.ch.

rationality principle, as discussed here, cannot be used for prescriptive purposes. Thus, we only discuss the positive variant of this principle.¹

When the content of the rationality assumption is discussed, one crucial point in the debate is whether it is used in a purely instrumental (formal) sense as ‘Zweckrationalität’ (in the sense of MAX WEBER) or whether it has some material content, that is, whether only the means are (or should be) rationally chosen in order to reach certain goals, whatever these goals are, or whether the goals themselves should be rationally justified. HABERMAS (1964), for example, accused ALBERT (1964) in the 1960s, in the famous Positivist Dispute in German Sociology, of a “positivistically halved rationalism”, because he insisted on the classical dichotomy between facts and standards, i.e. that it is not possible to provide value judgements with the same scientific justification (status) as statements about facts, which HABERMAS, with his consensus theory of truth, disputed.

Whenever such a substantive version of rationality is supposed, along with the prescriptive character of this assumption might come an ethical one: if some goals are justified and not striving for these goals would be irrational, such behaviour might not only be assessed as being imprudent but should also be avoided for moral reasons. Positive Economics accepts the dichotomy between facts and values, an idea that dates back to DAVID HUME.²

Following the classical model “that deliberation is always about means, never about ends” (SEARLE, 2001, p. 5), in Economics, a formal rationality assumption is usually employed. The philosophical idea behind this, coming from the period of the Enlightenment, is the autonomy of the person or, in economic terms, ‘consumer sovereignty’: individuals are free to choose their own values; nobody has the right to tell them what their preference function has to be.³ However, even then, there is considerable leeway for different concepts of rationality. They concern in particular (i) the content (and structure) of the individual utility function and (ii) the information used by the individual and how it is used by her/him. While the latter is mainly a subject of the concept of ‘bounded rationality’ as developed by SIMON (1955) and HEINER (1983), the former mainly centres on the notion of self-interest.

1 For the descriptive and prescriptive use of the rationality assumption see also SEN (1987).

2 On the role of the principle of value freedom in Economics see, for example, KIRCHGÄSSNER (2006).

3 This does not, of course, imply that nobody has the right to tell autonomous individuals how to behave properly; it is undisputed in Economics that the community has (up to a certain extent) the right to set rules for social living together and, in this way, to restrict the behavioural leeway of individuals.

Every strong version of the rationality principle has been empirically rejected. This result of modern behavioural economics becomes especially obvious if empirical results of game theory are taken into account.⁴ Laboratory experiments show that people behave according to the predictions of neither non-co-operative nor co-operative game theory; in particular, they deviate from the assumption of self-interest that is applied in formal models. Thus, not only the information problem is tackled today by behavioural economics, but also the motivation of acting individuals.

The empirical rejection of rationality concepts goes even further. Even in situations that largely favour the traditional notion of rational behaviour, people often (and consistently) behave differently. Starting with the Allais paradox of 1953,⁵ many anomalies of individual behaviour have been discovered, and some of them relate to the behaviour on financial markets: institutions that, at least until recently, were usually assumed to punish non-rational behaviour quite heavily. Nevertheless, there is evidence, for example, of Monday effects, excess volatility and the winner's curse.⁶ Moreover, tests of whether individuals have rational expectations in the sense of MUTH (1961) usually reject this assumption.

This does not necessarily imply, however, that such versions of the rationality principle should generally not be applied in economic analyses. In order to reject a theory it is necessary not only for it to be falsified but also for a 'better' theory to be available for the purpose at hand, a theory that has more explanatory power. Moreover, in particular with respect to the rational expectations assumption, it has to be taken into account that this is the most parsimonious possibility to construct macroeconomic models in such a way that they do not imply that individuals can be fooled consistently and/or in the long run in the same way. Despite the fact that the rational expectations hypothesis has often been rejected in empirical research, with the existence of money illusion being one of the reasons for these rejections,⁷ it hardly makes sense to evaluate different possibilities of economic policies under the assumption that money illusion exists in the long run, which actually implies that individuals do not learn and can be fooled even in the long run.

The topic of this paper is, however, the weakest possible version of the rationality principle. Contrary to the strong versions, this principle cannot be rejected

4 See for this, for example, SELTEN (1990).

5 For the history of the Allais paradox see SCHOEMAKER (1982, pp. 541 ff.).

6 See, for example, the corresponding chapters in THALER (1992).

7 See, for example, FEHR und TYRAN (2001).

by empirical evidence; it is more a methodological advice than an empirical statement.⁸ Nevertheless, it is, as we will see, one basic and necessary ingredient of any ‘understanding’ social science (‘verstehende’ Sozialwissenschaft) in the sense of MAX WEBER. Social sciences can manage without it, but not understanding ones in this sense.

Other basic ingredients of such an understanding social science are assumptions about the objectives of the individuals, about their intentions and about the information they have at hand and also actually use. Together, these assumptions form the economic model of behaviour. All this is based, of course, on the concept of methodological individualism, that is, on the assumption that individuals, human persons and not some collective agents, are those who act, even if they perform ‘common actions’. This requires a micro-foundation for economic theories, but also for other theories of society. Other concepts of social sciences that do not include such a micro-foundation, of course, do not have to employ any (weak) rationality assumption.

In the following, first the weak rationality principle and its role in economic theorising are discussed (*Section 2*). This section also discusses why it makes sense to use a micro-foundation and, therefore, to employ the rationality assumption in economic models. Then, we discuss whether the anomalies of individual behaviour as highlighted in modern behavioural economics do impair the applicability of the weak rationality principle (*Section 3*). It will be shown that this is not the case. Finally, we conclude with some remarks on handling the problems of ‘free will’ as well as ‘weakness of the will’ in the economic approach (*Section 4*).

2. The Rationality Principle

In economic analyses, the individual’s decision situation is essentially described by two elements: preferences and restrictions. Both elements are strictly differentiated in economic analyses.⁹ In a given situation, the restrictions limit the individual’s leeway for action; to these restrictions belong, besides others, the income of the individual, the market prices of goods, the legal frame of his actions but also the (expected) reactions of other individuals. Within this leeway, various

8 See for this also BRENNAN (2007, p. 121), who “argued for a conception of the ‘rational actor approach’ as providing not a theory of action so much as a grammar of argument about it”.

9 The economic approach differs in this respect from other approaches in the social sciences that do not make this distinction at all or at least not so strictly, for example, traditional sociology.

alternatives for acting are available, from which the individual can choose. It is unnecessary for the individual to know all the alternatives. Generally, he only knows part of his choices and this knowledge is often quite limited, so that he is only aware of some of their consequences. Before taking a decision, he must, therefore, evaluate these alternatives; he has to build up (conditional) expectations or forecasts.¹⁰ Nearly always one of the alternatives is to postpone the decision and to search for additional information in order to increase the knowledge about possible actions and their consequences. The preferences are derived from the intentions of the individual, they reflect the individual's ideas of value as they have been developed during the process of socialisation and they are principally independent of the actual possibilities of action. According to these preferences, the individual assesses the various alternatives at his disposal, he weighs up the pros and cons, the costs and benefits of the alternatives, against each other and finally chooses that (those) alternative(s) that come(s) closest to his preferences or that promise(s) to bring about the maximum net benefit.¹¹ Thus, in this model, human behaviour is interpreted as a rational choice from the available alternatives by the individual or – to speak in the language of economics – as ‘utility maximisation under constraints with uncertainty’. It is essentially a consequentialist approach: individuals weigh the alternatives at hand according to their expected consequences.¹²

Two issues are important when considering an individual's decision within the framework of the economic model of behaviour: the independence of the decision and the rationality of the decision. Independence of a decision means that the individuals act according to their own preferences (and not according to the preferences of others). Of course, they can take into account the interests of others in their preferences; in extreme cases they can be envious or malevolent, but also altruistic and benevolent. The exact content of the utility function is open.¹³

The second point is the rationality of the decision. In this context, rationality does not mean that the individual chooses the optimum way of acting at every

10 ESSER (1996) denoted this the “definition of the situation” that precedes every action.

11 RAWLS, who called this concept “the standard one familiar in social theory”, remarked that “in the usual way, a rational person is thought to have a coherent set of preferences between the options open to him. He ranks these options according to how well they further his purposes; he follows the plan which will satisfy more of his desires rather than less, and which has the greater chance of being successfully executed” (1971, p. 143).

12 See for this also SEN and WILLIAMS (1982). A detailed presentation of the economic model of behaviour is given by KIRCHGÄSSNER (2008).

13 In the terminology of SEN (1985, p. 347; 2007, p. 17) “self-goal choice” is assumed but – contrary to his belief – this does not rule out any kind of behaviour or motivation.

moment, that he goes through the world like a walking computer that always finds out the best of all the available alternatives in a flash. This distorted picture of the ‘homo oeconomicus’, which up to now is still found in many (text)books of microeconomics and which has rightly been criticised again and again, is not in line with the modern interpretations of the economic model of behaviour.¹⁴ Rationality in this model simply means that the individual, following his intentions, is principally in a position to assess and evaluate his action range and then to act accordingly.¹⁵ It has to be taken into account, however, that the individual has to make his decision without being fully informed and that the search for additional information is costly. He also often has to decide under time pressure. The individual will be especially willing to accept costs for additional information once he realises a relevant change of his action leeway and, therefore, has to assess and evaluate his alternatives once again. A rational individual reacts to such a change ‘systematically’, i.e. neither by chance nor randomly, but also not strongly traditionally insofar as he sticks strictly to the given rules, independent of the concrete situation.¹⁶

From this it follows that the behaviour of the individual can systematically be influenced by providing incentives, which in most cases result from changes in the individual’s action leeway (his restrictions). This makes predictions of behavioural changes as a reaction to changes in his action leeway, and, therefore, economic policy, possible. Moreover, in this concept, the philosophically meaningful and often-discussed distinction between human behaviour and human action disappears: the behaviour of individuals is explained by assuming that they act rationally.¹⁷

WEBER adopted a similar if not the same position when he wrote in 1922 about social behaviour on the first page of his *Economics and Society*: “It will be called human ‘behaviour’ only insofar as the person or persons involved engage in some

14 For criticism of this concept of rationality see also ARROW (1986).

15 WATKINS gave a similar definition of the ‘principle of rationality’, which he, however, called “just rough and provisional”: “An individual is placed in a certain objective problem-situation. He has certain aims (wants, preferences) or perhaps a single aim, and he makes a factual appraisal (which may be a misappraisal) of his problem-situation. The rationality principle says that he will act in a way that is ‘appropriate’ to his aim(s) and situational appraisal” (1970, p. 172). He explicitly ascribed the term ‘appropriate’ to POPPER (1967).

16 For the discussion of such kinds of ‘irrational’ behaviour see BECKER (1962) as well as, referring to him, ELSTER (1979, p. 137 ff.).

17 See also KIRCHGÄSSNER (1985) for this. This position is, of course, not uncontested. A different view is especially taken by those authors who combine the term ‘action’ with a moral demand, for example, PATZAK (1984).

subjectively meaningful action. Such behaviour may be mental or external; it may consist in action or omission to act. The term ‘social behaviour’ will be reserved for activities whose intent is related by the individuals involved to the conduct of others and is oriented accordingly” (1922, p. 1).

In other words, within the framework of the economic model of behaviour, individuals are supposed to adapt to changed environmental conditions according to their objectives, or preferences, in a systematic and therefore predictable manner. Such changes can result both from the actions of other individuals, for example, by political measures, as well as through changes in so-called ‘natural’ conditions. This was formulated as a principle by KLIEMT in the following way: “Every intentional human behaviour is to be explained as individual adaptive behaviour guided by preferences” (1984, p. 17).

According to the logic of science, this ‘weak principle of rationality’ might, as a basis for the economic model of behaviour, be of similar importance for the social sciences as the ‘principle of causality’ for the natural sciences.¹⁸ In the same way as in natural sciences talking about natural laws does not become possible before accepting the principle of causality, in the social sciences the understanding of human actions is impossible if the distinction between preferences and restrictions, or purposes and means, embedded in the economic model of behaviour is not accepted and if it is not presupposed that individuals use the means at their disposal in a subjectively rational way to reach their objectives. Of course, an important difference also exists between the two principles: whereas, as mentioned above, the principle of rationality can also be and is often comprehended in a normative sense, trivially this does not apply to the principle of causality. So far, the analogy between the two principles is limited.

It is interesting to note that not only the ‘new economic history’, as represented, for example, by NORTH (1981), but also the traditional ‘understanding’ branch of history applies exactly this ‘economic’ model of behaviour and, therefore, also the underlying rationality principle, although at first sight historical analysis and theoretical economics seem to be methodologically a great distance apart. That there is a large discrepancy might seem to be obvious if we think of the battle of methods (‘Methodenstreit’), started by CARL MENGER in 1883 between his, Austrian, theoretical school and GUSTAV SCHMOLLER’s historical school of political

18 See for this also POPPER (1967), TIETZEL (1981, p. 131 ff.), ABEL (1983, p. 133 ff.) and, for a somewhat different position, LATSIS (1983). Of course, the principle of causality can also be understood differently. (As to the importance of the principle of causality see, for example, STEGMÜLLER (1960) as well as, in particular, BUNGE (1959)).

economy.¹⁹ However, from a methodological perspective, those two approaches are much closer than one might think. In this context, for example, WATKINS wrote about the principle of rationality: “But the principle can also be cast in the form of a methodological rule that enjoins historians and other investigators of human behaviour, not necessarily to accept the principle *qua* factual postulate as true, but to proceed on the supposition that it is true. In this last form ... it says, first of all, that to provide a conjectural explanation for a past action is to postulate a decision-scheme which has a practical conclusion of which that action could be the natural outcome” (1970, p. 209).

Regarding it this way, the weak rationality principle is not an empirical statement; it can be neither verified nor falsified. Whenever an action of an individual is observed, it might be rational in the sense described above, even if we (at the moment) do not understand it. Thus, it cannot be falsified. On the other hand, even if we did have a ‘rational’ explanation for every observed action, we could never be sure that such an explanation is really feasible for all human actions. Thus, this principle cannot be verified either. Statements that can be neither verified nor falsified have no empirical content and are, therefore, not scientific ones in a strict sense. They are metaphysical statements. This also holds for the weak rationality principle. Its purpose is not to tell us something about the reality; it is a heuristic principle that gives us methodological guidance for the way we are conducting research in the social sciences. It tells us that whenever we want to explain human behaviour, we should try to explain it as the rational actions of individuals or as the result of rational decisions, respectively.²⁰

This is somewhat contrary to POPPER’s view of the rationality principle. On the one hand, there is at best a minor disagreement about the content of the principle and its importance for the social sciences. He classified the economic approach as “*a purely objective method* in the social sciences which may well be called the method of *objective* understanding, or situational logic” (1962, p. 199). In his *Objective Knowledge* he called this procedure “situational analysis” (1972, p. 178). He defined the rationality principle as “the principle of acting appropriately to the situation” (p. 359). On the other hand, there is some disagreement about the status of this principle. Though he acknowledged that it is “clearly an

19 See for this, for example, SCHUMPETER (1954, p. 814f.).

20 See also VANBERG (2004, p. 3): “Interpreted as a heuristic principle, the rationality principle tells us how we should go about explaining purposeful human action. It suggests that we ought to explain such action in terms of the actor’s purposes and beliefs, and that we should do so under the presumption that the actor’s purposes and beliefs are consistent at the moment of choice.”

almost empty principle” (1967, p. 359), he believed that it has some empirical content and, therefore, can (in principle) be tested. Mentioning *prima facie* evidence, he claimed that “the rationality principle is false” (1967, p. 361). Nevertheless, he also acknowledged its heuristic value.²¹

If this behavioural model, as it is supposed here, is seen as a prerequisite for understanding human action, the difference between ‘explanation’ and ‘understanding’ disappears: I can only understand human action if I can explain it by means of such a model of rational behaviour. MAX WEBER, to whom is usually attributed the ‘understanding’ branch of the social sciences, put this into similar words. He looked at sociology as “that science which aims at the interpretative understanding of social behaviour in order to gain an explanation of its causes, its course, and its effects” (1922, p. 1).²² This also corresponds to our ordinary language usage of ‘understanding’ of human behaviour. Here, we also think in categories of motives (preferences), means (restrictions) and limited information.²³

An individual might act rationally in this sense even if he does – according to our everyday perceptions – rather crazy things. Someone, for example, who is schizophrenic and believes that he is another person and/or in a totally different situation might be ‘rational’ in this sense, given his (totally wrong) perception of the reality. Moreover, we can only understand him if we – underlying his perception – ask why he behaved in this and that way. As soon as we ask “why” somebody behaves in a certain way, we usually assume that he believed there were ‘reasons’ for his behaviour and, correspondingly, acted according to the weak rationality principle.

There is only one ‘strong’ assumption in this concept: consistency at the moment of choice. Preferences are assumed to be complete, but only insofar as the individual is assumed to be able to evaluate principally any alternative that he takes into account as a possibility, and the ordering of these possibilities is assumed to be transitive. Without this assumption, we cannot ‘understand’ what an individual does. However, even this assumption is not testable if we do not assume that the preferences are more or less constant over time. Actions at different points in time that seem to be inconsistent might be consistent if the individual learns, and/or has only limited information processing capacity or changes his evaluations for whatever reason. Moreover, learning, and/or forgetting, might

21 For a discussion and a critique of POPPER’s conception of the social sciences see VANBERG (1975, pp. 109 ff.), SCHMID (1979, 1979a) as well as HANDS (1985).

22 Regarding MAX WEBER’s position see, for example, ANGEHRN (1983).

23 For a discussion of the term ‘understanding’ in social sciences see, for example, STEGMÜLLER (1969, p. 360 ff.), BÜHLER (1987) as well as the contributions by SCHURZ (1988).

change the individual's behaviour over time in a way that seems to be inconsistent. Thus, consistency over time is a necessary assumption when revealed preference analysis is to be applied, but it is not a necessary component of the rationality assumption.²⁴

Applying the economic model of behaviour is, of course, not the only way by far in which social sciences can proceed. Methodological individualism and the rationality assumption are not necessary elements of every social science. Correspondingly, a micro-foundation is not a necessary ingredient either.²⁵ Macro-theories can be developed that discuss, for example, the properties of social systems and the relations between those properties that do not need a micro-foundation and, therefore, also do not need to apply the weak rationality principle. This also holds for economics, in particular because economics is hardly interested in the behaviour of single individuals, but in the behaviour of so-called 'aggregates', such as consumers, entrepreneurs or voters. It is not the behaviour of a certain single individual that is interesting, but the 'typical' behaviour: regularities in the behaviour of all or at least the majority of the individuals in the respective group.²⁶ Here, the micro-theory offers (only) the basis in order to be able to explain the macro-phenomena.

This is not a contradiction, as it might seem at first glance. If by change of a certain macro-variable the conditions for the actions of all the individuals in a certain group are influenced in a similar way, it is to be expected that their reaction will not in every single case, but on average, show that regularity that can be explained by the individual decision calculus. Thus, a rise in petrol prices will, for example, not induce every car driver – *ceteris paribus* – to save petrol. For the economic way of reasoning it is, however, only relevant that *on average* consumers react with savings so that the rise in prices leads to a reduction in the total demanded quantity. This behaviour, which actually could be observed after several large increases in petrol prices in recent decades, can – by using some additional 'weak' assumptions – be derived for the 'typical' consumer from the individual optimality calculus of the theory of consumer behaviour.

24 See, for example, SUDGEN (1985).

25 For a discussion of the two sociologies, the one with and the other without such a micro-foundation, see VANBERG (1975).

26 Correspondingly, HICKS wrote when dealing with the law of demand: "In all our discussions so far, we have been concerned with the behaviour of a single individual. But economics is not, in the end, much interested in the behaviour of single individuals. Its concern is with the behaviour of groups. A study of individual demand is only a means to the study of market demand" (1939, p. 34). – See for this also v. HAYEK (1952, p. 48 ff.) as well as POPPER (1967, p. 3).

Whether a micro-foundation is necessary for economic theories was, for example, discussed in the 1970s.²⁷ Macro-relations based on arguments of plausibility were accepted as long as they seemed to be empirically valid; using econometric methods, the latter had to be ‘proven’ by statistical significance. Economic policy was performed by relying on the existence and stability of such relations.

The most famous example of such a relation is probably the well-known modified Phillips curve, which contains a long-term trade-off between inflation and unemployment. It was first ‘discovered’ by PHILLIPS (1958) as a relation between unemployment and real-wage development. Later on, it was expanded into a relation between inflation and unemployment. In 1960 it was called the “menu of choice” of economic policy by SAMUELSON and SOLOW (1960, p. 192). Such Phillips curves, named after their ‘discoverer’, were econometrically estimated for quite a number of countries and, in the 1970s, contained in all the large macro-econometric models. On the basis of this empirical evidence it was believed that the unemployment rate could be permanently reduced through a once and for all increase in the inflation rate.²⁸

Such macroeconomic relations – respectively the macroeconometric models consisting of them – can in many cases be helpful instruments for forecasts. This can change, however, as LUCAS (1976) pointed out in his famous critique, if one tries to simulate the effects of different economic policies with such models in order to put into practice those economic policy measures that should lead to the ‘best’ result. Then it might happen that a macroeconomic relation, which ought to be exploited for economic policy purposes, disappears. If, for example, a politician wants to make use of the (statistically validated) long-term relation between inflation and unemployment by increasing inflation through a policy of easy money in order to reduce unemployment, this might perhaps be successful in the short run. In the medium or long run, however, inflation will increase while unemployment will again reach its former level. As the 1970s taught us, in the end, it might even be somewhat higher than before.

The political, or so to say ‘experimental’, demonstration of the Phillips curve in the 1970s and 1980s failed: prices continued to rise along with unemployment. The reason for this failure was the fact that this policy relied on money illusion of the economic agents. However, such a policy can only be successful in the long run if economic agents can be fooled systematically and permanently. This does not seem to be very plausible. It is also hardly compatible with any concept

27 See for this, for example, RAMSER (1987, p. 8 ff.), as well as MACHLUP (1963) and SCHLICHT (1977).

28 For this and the respective criticism see SANTAMERO and SEATER (1978) as well as the discussion between KING and WATSON (1994) and EVANS (1994).

of rationality. Consequently, the necessity of a microeconomic basis, and therefore of the application of the weak rationality principle, is accepted today quite generally in economics.

3. The Rationality Assumption and the Results of Behavioural Economics

As mentioned above, the rationality principle can be neither verified nor falsified: it has a metaphysical character but can be – and often is – a useful heuristic device. Nevertheless, its empirical content is zero. A priori, this might be no major problem if, for example, a historian tries to explain a historic episode, but even then he has to make additional assumptions about the motivations of the acting persons and about the information available to them. If these assumptions are wrong, his explanation might fail.

There is, however, a much larger problem if we want to derive general hypotheses about human behaviour and, in particular, as is usual in Economics, about the social outcomes of individual behaviour. In this case, like the historian, we also have to make additional assumptions about the motivations of the acting persons as well as about their informational status. The stricter these assumptions are, the larger is the content of our theory, but the larger is also the probability that our theory will be rejected, or at least that some hypotheses derived from this theory will be rejected, by the resistance of reality. Whenever this is the case, it is not the weak rationality assumption that is rejected; it is at least one of the additional assumptions that is responsible for this failure.

When performing such tests, for example with experiments, it is often stated that people do not behave rationally and/or that the economic model of behaviour, the ‘homo oeconomicus’, is rejected. Some people even claim that “The homo oeconomicus is dead”.²⁹ However, this is a categorical mistake. What is rejected in these experiments is mostly the assumption of pure self-interest, that is, the motivational assumption usually employed in economic analyses; it is not the rationality assumption. Despite the fact that this is often done, even by such famous economists as SEN, one should not mix up these two.³⁰ It is obvious, that,

29 See, for example, HÄRING, “Der Homo Oeconomicus ist tot”, *Financial Times Deutschland* of 14 March 2001, http://ockenfels.uni-koeln.de/fileadmin/wiso_fak/stawi-ockenfels/pdf/Presse/Der_Homo_oeconomicus_ist_tot.pdf (26/06/12).

30 See, for example, HAUSMAN and MCPHERSON (1996, p. 2), who explicitly mentioned that in positive economics the assumption of self-interest is added to the theory of rationality. On the

in addition to self-interest, people have other motives as well. Whether the parsimonious assumption of self-interest is appropriate to analyse a specific situation depends on the concrete circumstances and is not a question of whether pure self-interest – in the narrow sense – is the only motivation people have. It is even less a question of rationality.

This is obvious if we consider the famous 1999 paper by FEHR and SCHMIDT. To derive a theory that is consistent with their experimental results, they adjusted the utility function of individuals by including inequality aversion and then applied the full maximisation procedure. Thus, in the framework of the traditional approach, understanding rational behaviour as utility maximisation under constraints, their procedure is nothing other than applying exactly this approach with a slight change in the utility function. From a methodological point of view this is only a slight change, even if it has considerable consequences for the explanatory power of the model.

There are at least three phenomena that are empirically well-established, but not consistent with the assumption of narrow self-interest as defined above: altruism, reciprocity and commitment. If, according to SEN (1977, p. 95), ‘one way of defining commitment is in terms of a person choosing an act that he believes will yield a lower level of personal welfare to him than an alternative that is also available to him’, in all three cases individuals impose costs upon themselves that do not have a direct positive affect on their individual welfare. In the case of altruism it has a direct positive impact on other individuals. In the case of reciprocity this might be positive or negative, depending on the previous action of other individuals; the really interesting case is the one of negative reciprocity in which the individual incurs costs in order to punish another individual.³¹ Following the above definition in the case of commitment this is open: a person might be committed to a behaviour that benefits somebody else, a behaviour that hurts other people, but also to a behaviour that does not affect other people at all. Accordingly, altruism and reciprocity might be seen as special cases of commitment. The main difference is the sign of the corresponding element in the individual’s utility function.³²

other hand, SEN (2005, p. 5), for example, contributed to the confusion when he wrote that the focus of rational choice theory is “on characterising rationality of choice as maximisation of self-interest”.

31 Positive reciprocity, such as in the case of tit for tat, is in the (long-run) self-interest of the individual and, therefore, not at odds with the usual assumptions employed in economic models.

32 An analogous argument with respect to commitment can be made for value rationality in the sense of MAX WEBER.

It has been shown how altruism, reciprocity or moral behaviour can be incorporated into economic models of utility maximisation subject to constraints.³³ There is also no problem with formalising commitment along these lines. The only problem is how to specify the utility function. This does not touch the (weak) rationality assumption at all.³⁴ There is, for example, no conflict between morality and rationality. Moreover, the ‘rational fool’ of SEN (1977) can be considered as being a fool, not because of his irrationality but because of the special structure of his utility function (with a rather high discount rate).

One might question whether it is appropriate to integrate these kinds of behaviour into the (standard) economic model, in order to make “more room for other regarding preferences in the formulation of rational choice” (SEN, 1994, p. 389). The fact that moral phenomena can be accommodated within utility theory does not yet imply that it is appropriate to analyse the phenomena at hand in this way.³⁵ However, this is a question of appropriateness and not of possibility. The use of models based on this ‘economic’ framework allows us to deduce testable hypotheses about which conditions are favourable for other-regarding behaviour.³⁶ The theoretical and empirical literature about voter participation, which, since RIKER and ORDESHOOK (1968), has included moral behaviour in a standard economic model, can serve as a good example.³⁷ Moreover, it remains open what an alternative representation in a formal model should look like. Alternatively, should one totally dismiss formal models in this context? Finally, even if (standard) utility theory is not applied, it remains that acting individuals have objectives and a perception of the alternatives that are available to them, and that we have to apply the weak rationality principle in order to understand their behaviour.

As mentioned above, in other experiments, strong models of rationality have been rejected, as was demonstrated first at the famous Paris conference in May 1952 when SAVAGE tried to defend his independence axiom. This axiom

33 For altruism see, for example, ANDREONI (1988, 1989, 1990), for reciprocity FEHR and GÄCHTER (1998, 2000) and for moral behaviour KIRCHGÄSSNER (2002, 2010).

34 See for this also BRENNAN (2007, p. 115), who stated that, thinking “of the ‘rational actor approach’ as an approach rather than a theory; and, second, that if one thinks of rationality in these terms, then SEN falls firmly within the rational actor school”.

35 HAUSMAN und McPHERSON (1996, p. 53) argue “that some moral phenomena cannot be accommodated within utility theory”.

36 See, for example, G. KIRCHGÄSSNER (1992).

37 The fact that GREEN and SHAPIRO (1994) used voter participation as one of their examples of the empirical failure of the rational choice approach in politics does not contradict this. Assuming self-interest, they only considered the instrumental value of participation and are thus far correct with their critique. However, this is just an example of the problems arising when rationality and self-interest assumptions are not distinguished.

is implicitly included in the von Neumann–Morgenstern concept of expected utility maximisation. ALLAIS was able to demonstrate that even SAVAGE himself violated his own axiom.³⁸ In this and similar cases, the concrete model of rationality that has been presupposed is rejected. Nonetheless, this is definitely not a rejection of the weak rationality principle as well; it is perhaps the strongest version of the rationality principle that is rejected.

In these classical experiments, people are usually fully informed. In other experiments and, in particular, in real decision situations, people are rarely fully informed; usually they have only very limited information, be it about the alternatives available to them or about the possible consequences. Here, the concept of bounded or procedural rationality, as first developed by SIMON (1955), enters.³⁹ In such situations, the behaviour might differ substantially from what would be expected under ‘full’ or ‘substantive’ rationality. Even then, though, it is not only the weak rationality that is applied. It is the economic model with its basic assumptions: that people have intentions and a – more or less correct or incorrect – picture of the reality, and, based on this, their acting possibilities, and, whenever they are following a satisficing strategy, that they also follow the weak rationality principle by choosing their most favourable alternative, given their information and their expectations.

The situation might seem to be quite different if we observe framing effects, as in the famous experiments by TVERSKY and KAHNEMAN (1981, 1987).⁴⁰ Here, the consistency requirement discussed above might be violated. However, this does not need to be the case. If two different frames suggest two different realities, reasonable people might show two different reactions. Thus, it is a problem of information and information processing, but not of rationality, and we are, therefore, back in a situation of bounded rationality. There is no need to speak of irrationality. This is obvious if we take into account that this inconsistency vanishes as soon as individuals become aware of the substantive equivalence of the two frames.

A stronger argument against the concept of weak rationality might be derived from preference reversal experiments. In these experiments, the transitivity condition mentioned above seems to have been violated.⁴¹ However, what has really

38 See ALLAIS (1953) and, for the discussion at this conference, FISHBURN (1987).

39 See for this also SIMON (1978, 1979), SELTEN (1990), CONLISK (1996) as well as RUBINSTEIN (1998).

40 See for this also STOCKÉ (2002).

41 See for this LICHTENSTEIN and SLOVIC (1971), POMMEREHNE, SCHNEIDER and ZWEIFEL (1982), GUALA (2000) as well as the review of TVERSKY and THALER (1990). That preferences might show “consistent und predictable intransitivities” has already been shown by TVERSKY (1969, p. 31).

been violated here is the assumption of consistence over time. It might again be a problem of different frames. If, as TVERSKY (1969) stated, such behaviour is consistent with regret theory, it is obvious that it is also again a problem of the traditional concept of rationality and not one of the weak rationality principle: his arguments are based on a general model of choice.

It might be objected that in reality, the behaviour of individuals is characterised less by rational or bounded rational decisions, but more by their adherence to legal or social norms. This role model of human behaviour has, for example, been presented by DAHRENDORF (1958) in his picture of the 'homo sociologicus'. This sociological model can be seen in the tradition of a non-individualistic sociology or social sciences going back to DURKHEIM's *Rules of Sociological Method* published in 1895. Because people only rarely take very deliberate decisions but often make decisions intuitively, this model is often contrasted with the economic model of behaviour under the pretension of being a 'more realistic one'.

However, as long as one is not strictly following such rules, i.e. independent of the concrete situation, rule-following behaviour is totally compatible with the weak rationality principle. In DAHRENDORF's (1958) model, individuals react to sanctions, that is, to incentives, in the same way as in the economic approach. Moreover, VANBERG (1975) correctly classified DAHRENDORF's approach as an individualistic social theory, as one that is based on the general economic model of behaviour. Finally, KÖNIG (1961, pp. 36 ff.) pointed out the fact that it might be possible to reinterpret the most famous work of DURKHEIM, his *Suicide* of 1897, in an individualistic manner. DURKHEIM's main intention was to show that social facts are not directly deduced from individuals' intentions, but are often unintended consequences caused by social conditions. In this, he is quite in line with ADAM SMITH and the economic tradition. Thus, notwithstanding the fact that there are sociological theories of society that it might be very difficult or even impossible to reconstruct in an individualistic manner, the sociological role theory does not contradict the weak rationality principle.

The final question to be discussed here is whether the new emerging field of 'Neuroeconomics' causes doubts about the validity of the weak rationality principle.⁴² On the one hand, this can obviously not be the case as this principle is not a statement with empirical content. On the other hand, some results of this new field might even support rather strong versions of the economic model. By scanning the brain of subjects, DE QUERAIN et al. (2004), for example, showed that punishment actions, which incur costs for those who are punishing, activate

42 On recent developments in Neuroeconomics see, for example, FEHR and RANGEL (2011) as well as the contributions in the *Journal of Economic Methodology* 17/2, June 2010.

the same areas of the brain that are activated by actions that generate (in the traditional sense) satisfaction like the consumption of goods. This allows us to introduce psychological effects into economic models in a non-tautological, empirically testable way. Thus, the results of some neuroeconomic investigations might even rehabilitate the traditional model of the homo oeconomicus in some respects. People seem to try to obtain as much satisfaction as possible and such satisfaction may also result from actions that contradict the narrow version of self-interest: they are costly but do not allow for additional consumption possibilities. However, this is no contradiction at all to the weak rationality principle.

Finally, what has been a heuristic principle so far might perhaps become a testable hypothesis with the use of neuroeconomic methods. Accordingly, Neuroeconomics might lead to a microfoundation of the micro-foundation of Economics (and other theories of society).⁴³ Whether this will really be the case and which consequences this will have for the explanation of macro phenomena and, in particular, for the use of economic knowledge in economic policy remains unresolved.

4. Summary and Concluding Remarks

The weak rationality principle is a heuristic principle that guides any ‘understanding’ social science in the Weberian sense. It has no empirical content and can, therefore, not be rejected for empirical reasons, either in experiments or in field studies. It could, however, be rejected for practical reasons if, for example, following another heuristic principle would lead to ‘better’ insights into social relations and/or to more fruitful theories of the society. Whether this is the case or not is not the topic of this paper, even if it is difficult to imagine an alternative that would generally outperform the general rational choice model. This does not imply that other models cannot have advantages in explaining some phenomena.

Empirical tests are principally (and often rather easily) possible if assumptions are added that specify the preferences and/or the available information in more detail, that is, if additional hypotheses are added to the weak rationality principle. This holds, for example, for the various tests of the rational expectations

43 According to FEHR and RANGEL (2011, p. 4), referring to WILSON (1999), “the ultimate goal [of neuroeconomics] is to produce detailed computational and neurobiological accounts of the decisionmaking process that can serve as a common foundation for understanding human behavior across the natural and social sciences”. Thus, a micro-foundation of not only microeconomics, but of all sciences that investigate human decision behaviour, is sought.

hypothesis.⁴⁴ Then, however, ‘combined hypotheses’ are tested, and it is hardly possible to decide whether (in case of a failure) the assumption of rationality and therewith the basic behavioural model or (one of) the additional special assumption(s) has to be rejected.

One advantage of applying the (weak) rationality principle is that it makes two seemingly contradictory views compatible: the assumption of free will and the predictability of human behaviour. To predict the behaviour of an individual it does not matter whether he follows a deterministic law or rationally decides to behave in a certain way.⁴⁵ Ex post we can, of course, try to understand any behaviour. This, however, does not necessarily help to predict future behaviour. In order to achieve the latter, we have to construct models that represent the typical situation of the individual, the ‘situational logic’ in POPPER’s terms (1972, p. 102). Because individual decisions also depend on many factors that cannot be explicitly included in our micro-models, the rationality principle is ex ante more useful for predicting aggregate compared with individual behaviour, provided that the individual ‘deviations’ of the solution proposed by the model cancel each other out.⁴⁶

Finally, the weak rationality principle is also compatible with the idea of hierarchical preference orderings. Problems like the ‘weakness of the will’ and the observation that individuals deliberately restrict their future action leeway in order to reach certain goals cannot be handled within the traditional economic framework.⁴⁷ To do so, it has to be assumed that individuals have two kinds of preferences, which are ordered.⁴⁸ On the lower level are the preferences given in daily life and according to which the individual has to make decisions in concrete situations. The preferences on the upper level describe how the individual would like to see himself acting. Both preference orderings can be consistent, but contradictions can exist between the two orderings.⁴⁹ As long as no concrete actions are necessary, the individual might ‘plan’ according to his upper preferences, and he can try to influence those restrictions that later will be relevant in concrete situations, to make sure that despite following the lower preferences his actual behaviour will – as far as possible – be consistent with his upper preference ordering.

44 See, for example, KIRCHGÄSSNER (1993).

45 See for this LAGUEUX (2004, p. 32).

46 See for this also ZINTL (1989).

47 A whole collection of examples of self-restraint is given by SCHELLING (1984). Many of these examples are from everyday life.

48 See for this also SEN (1977), SCHELLING (1978, 1980), ELSTER (1986), as well as THALER and SHEFRIN (1981).

49 Such contradictions are especially obvious whenever addicts try to fight their addiction.

At first, the idea of two different and partly contradictory preference orderings might seem to be strange; it is unfamiliar at least for the economist who is used to assuming unambiguous and fixed preferences of individuals. This does not hold for the philosophical tradition. In this tradition, the idea can be traced back at least to ARISTOTLE and his *Nicomachean Ethics*, in which he spoke of two different parts of the soul whereby “the part of the soul that has a rational principle stands to the irrational part”, and he sees the relation between these two parts as “between ruler and ruled”.⁵⁰ FRANKFURT (1971) also argued that the existence of superior preferences and the possibility of using them for judging the subordinated ones is a precondition for seeing human beings as ‘persons’ with a free will because for this it is necessary for them to be conscious of their will and to be able to reflect on it. Exactly this happens when individuals are trying to bind themselves. This idea and even the terminology of ARISTOTLE were taken up, for example, by THALER and SHEFRIN (1981) in their Economic Model of Self-Control when they spoke of a planner and a doer who might have different preference orderings.

Thus, as its application in economics has shown, the weak rationality principle constitutes very powerful methodological guidance that can also be applied outside the economic area. Which additional assumptions have to be or should be made depends on the area of application. Whether it is fruitful in these applications mainly depends on the fruitfulness of the competing approaches available: those that are not within the framework of an ‘understanding’ social science. So far, I do not see important and powerful alternatives.

References

- ABEL, BODO (1983), *Grundlagen der Erklärung menschlichen Handelns*, Tübingen: Mohr (Siebeck).
- ALBERT, HANS (1964), „Der Mythos der totalen Vernunft: Dialektische Ansprüche im Lichte undialektischer Kritik“, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 16, pp. 225–256.
- ALLAIS, MAURICE (1953), «Le comportement de l’homme rationnel devant risque: Critique des postulats et axiomes de l’école Américaine», *Econometrica* 21, pp. 503–546; extended English version: “The Foundations of a Positive

50 ARISTOTLE, *Nicomachean Ethics*, 1138b.

- Theory of Choice Involving Risk and a Criticism of the Postulate and Axioms of the American School”, in: M. Allais and O. Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*, pp. 27–145, Dordrecht: Reidel 1979.
- ANDREONI, JAMES (1988), “Privately Produced Public Goods in a Large Economy: The Limits of Altruism”, *Journal of Public Economics*, 35, pp. 57–73.
- ANDREONI, JAMES (1989), “Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence”, *Journal of Political Economy*, 97, pp. 1447–1458.
- ANDREONI, JAMES (1990), “Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving”, *Economic Journal*, 100, pp. 464–477.
- ANGEHRN, EMIL (1983), „Handlungserklärung und Rationalität, Zur Methodologie Max Webers“, *Zeitschrift für philosophische Forschung*, 37, pp. 341–362.
- ARROW, KENNETH J. (1986), “Rationality of Self and Others in an Economic System”, *Journal of Business*, 59, pp. 385–399.
- BECKER, GARY S. (1962), “Irrational Behavior and Economic Theory”, *Journal of Political Economy*, 70, pp. 1–13.
- BRENNAN, GEOFFREY (2007), “The Grammar of Rationality”, in: F. Peter and H.B. Schmid (eds.), *Rationality and Commitment*, pp. 105–123, Oxford: Oxford University Press.
- BÜHLER, AXEL (1987), „Die Einheit der wissenschaftlichen Methode und Maximen des Verstehens“, *Zeitschrift für philosophische Forschung*, 41, pp. 633–644.
- BUNGE, MARIO (1959), *Causality: The Place of the Causal Principle in Modern Science*, Cambridge (Mass.): Harvard University Press.
- CONLISK, JOHN (1996), “Why Bounded Rationality?”, *Journal of Economic Literature*, 34, pp. 669–700.
- DAHRENDORF, RALF (1958), „Homo Sociologicus: Versuch zur Geschichte, Bedeutung und Kritik der Kategorie der sozialen Rolle“, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 10, pp. 178–208; English translation: “Homo Sociologicus: On the History, Significance, and Limits of the Category of Social Role”, in: R. Dahrendorf, *Essays in the Theory of Society*, pp. 19–87, Stanford: Stanford University Press 1968.
- DURKHEIM, EMILE (1895), *Les règles de la méthode sociologique*, Paris: Presses Universitaires de France; English translation: *The Rules of Sociological Method*, Glencoe (Ill.) : The Free Press, 1938.
- DURKHEIM, EMILE (1897), *Le suicide, Etude de sociologie*, Paris; English translation: *Suicide: A Study in Sociology*, New York : Free Press 1951.
- ELSTER, JON (1979), *Ulysses and the Sirens, Studies in Rationality and Irrationality*, Cambridge: Cambridge University Press.
- ELSTER, JON (ed.) (1986), *The Multiple Self*, Cambridge: Cambridge University Press.

- ESSER, HARTMUT (1996), „Die Definition der Situation“, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 48, pp. 1–34.
- EVANS, CHARLES L. (1994), “The Post-War U.S. Phillips Curve: A Comment”, *Carnegie-Rochester Conference Series on Public Policy*, 41, pp. 221–230.
- FEHR, ERNST, and SIMON GÄCHTER (1998), “Reciprocity and Economics: The Economic Implications of Homo Reciprocans”, *European Economic Review*, 42, pp. 845–859.
- FEHR, ERNST, and SIMON GÄCHTER (2000), “Fairness and Retaliation: The Economics of Reciprocity”, *Journal of Economic Perspectives*, 14(3), pp. 159–181.
- FEHR, ERNST, and ANTONIO RANGEL (2011), “Neuroeconomic Foundations of Economic Choice: Recent Advances”, *Journal of Economic Perspectives*, 25(4), pp. 3–30.
- FEHR, ERNST, and KLAUS M. SCHMIDT (1999), “A Theory of Fairness, Competition, and Cooperation”, *The Quarterly Journal of Economics*, 114, pp. 817–868.
- FEHR, ERNST, and JEAN-ROBERT TYRAN (2001), “Does Money Illusion Matter?”, *American Economic Review*, 91, pp. 1239–1262.
- FISHBURN, PETER C. (1987), “Reconsiderations in the Foundations of Decision Under Uncertainty”, *Economic Journal*, 97, pp. 825–841.
- FRANKFURT, HARRY G. (1971), “Freedom of the Will and the Concept of a Person”, *Journal of Philosophy*, 68, pp. 5–20.
- GUALA, FRANCESCO (2000), “Artefacts in Experimental Economics: Preference Reversals and the Becker-DeGroot-Marschak Mechanism”, *Economics and Philosophy*, 16, pp. 47–75.
- GREEN, DONALD P., and IAN SHAPIRO (1994), *Pathologies of Rational Choice: A Critique of Applications in Political Science*, New Haven: Yale University Press.
- HABERMAS, JÜRGEN (1964), „Gegen einen positivistisch halbierten Rationalismus: Erwiderung eines Pamphlets“, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 16, pp. 636–659.
- HANDS, DOUGLAS W. (1985), “Karl Popper and Economic Methodology”, *Economics and Philosophy*, 1, pp. 63–99.
- HAUSMAN, DANIEL M., and MICHAEL S. MCPHERSON (1996), *Economic Analysis and Moral Philosophy*, Cambridge (U.K.): Cambridge University Press.
- HAYEK, FRIEDRICH AUGUST VON (1952), *The Counter Revolution of Science: Studies in the Abuse of Reason*, Glencoe: The Free Press.
- HEINER, RONALD A. (1983), “The Origin of Predictable Behavior”, *American Economic Review*, 73, pp. 560–595.
- HICKS, JOHN R. (1939), *Value and Capital, An Inquiry into Some Fundamental Principles of Economic Theory*, Oxford: Oxford University Press. (2nd edition 1946.)

- KING, ROBERT G., and MARK W. WATSON (1994), "The Post-War U.S. Phillips Curve", *Carnegie-Rochester Conference Series on Public Policy*, 41, pp. 157–219.
- KIRCHGÄSSNER, GEBHARD (1985), „Rationales Verhalten und vernünftiges Handeln, ein Widerspruch?“, in: H. Milde and H.G. Monissen (eds.), *Rationale Wirtschaftspolitik in komplexen Gesellschaften*, pp. 29–41, Stuttgart: Kohlhammer.
- KIRCHGÄSSNER, GEBHARD (1992), "Towards a Theory of Low-Cost Decisions", *European Journal of Political Economy*, 8, pp. 305–320.
- KIRCHGÄSSNER, GEBHARD (1993), "Testing Weak Rationality of Forecasts with Different Time Horizons", *Journal of Forecasting*, 12, pp. 541–558.
- KIRCHGÄSSNER, GEBHARD (2002), "On the Role of Heroes in Political and Economic Processes", *Kyklos*, 55, 179–196.
- KIRCHGÄSSNER, GEBHARD (2006), „Wertfreiheit und Objektivität in den Wirtschaftswissenschaften: Mythos oder Realität“, in: G. Zecha (ed.), *Werte in den Wissenschaften: 100 Jahre nach Max Weber*, pp. 137–171, Tübingen: Mohr Siebeck.
- KIRCHGÄSSNER, GEBHARD (2008), *Homo Oeconomicus: The Economic Model of Behaviour and Its Applications to Economics and Other Social Sciences*, New York: Springer.
- KIRCHGÄSSNER, GEBHARD (2010), "On Minimal Morals", *European Journal of Political Economy*, 26, pp. 330–339.
- KLIEMT, HARTMUT (1984), „Nicht-explanative Funktionen eines ‚Homo oeconomicus‘ und Beschränkungen seiner explanativen Rolle“, in: M.J. HOLLER (ed.), *Homo oeconomicus II*, pp. 7–49, München.
- KÖNIG, RENÉ (1961), Einleitung (Introduction) to the German edition of E. Durkheim (1895): *Die Regeln der soziologischen Methode*, pp. 19–82, Neuwied: Luchterhand.
- LAGUEUX, MAURICE (2004), "The Forgotten Role of the Rationality Principle in Economics", *Journal of Economic Methodology*, 11, pp. 31–51.
- LATSIS, SPIRO J. (1983), "The Role and Status of the Rationality Principle in the Social Sciences", in: R.S. Cohen and M.W. Wartofsky (eds.), *Epistemology, Methodology and the Social Sciences*, pp. 123–151, Dordrecht/Boston: Reidel Publishing Company.
- LICHTENSTEIN, SARAH, and PAUL SLOVIC (1971), "Reversals of Preference Between Bids and Choices in Gambling Decisions", *Journal of Experimental Psychology*, 89, pp. 46–55.
- LUCAS, ROBERT E. (1976), "Econometric Policy Evaluation, A Critique", *Carnegie-Rochester Conference Series on Public Policy*, 1, pp. 19–46.

- MACHLUP, FRITZ (1963), "Micro- and Macroeconomics, Contested Boundaries and Claims of Superiority", in: F. Machlup, *Essays on Economic Semantics*, pp. 97–144, Prentice-Hall, Englewood Cliffs.
- MENGER, CARL (1883), *Untersuchungen über die Methode der Sozialwissenschaften und der politischen Ökonomie insbesondere*, Leipzig: Duncker und Humblot; English translation: *Problems of Economics and Sociology*, Urbana: University of Illinois Press, 1963.
- MUTH, JOHN E. (1961), "Rational Expectations and the Theory of Price Movements", *Econometrica*, 29, pp. 315–335.
- NEUMANN, JOHN VON, and OSKAR MORGENSTERN (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.
- NORTH, DOUGLASS C. (1981), *Structure and Change in Economic History*, New York: Norton.
- PATZAK, MELITTA (1984), „Rationalmodell und analytische Handlungstheorie“, in: M.J. HOLLER (ed.), *Homo oeconomicus III*, pp. 7–42, München.
- PHILLIPS, ALBAN W. (1958), "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom", 1861–1957, *Economica*, 25, pp. 283–299.
- POMMEREHNE, WERNER, FRIEDRICH SCHNEIDER and PETER ZWEIFEL (1982), "Economic Theory of Choice and the Preference Reversal Phenomenon, A Reexamination", *American Economic Review*, 72, pp. 569–574.
- POPPER, KARL R. (1962), „Die Logik der Sozialwissenschaften“, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 14, pp. 233–248.
- POPPER, KARL R. (1967), « La rationalité et le statut du principe de rationalité », in E.M. Classen (ed.), *Les fondements philosophiques des systèmes économiques: Textes de Jaques Rueff et essais rédigés en son honneur*, pp. 142–150, Paris: Payot; English translation: "The Rationality Principle", reprinted in: B.J. Caldwell (ed.), *The Philosophy and Methodology of Economics*, Volume III, pp. 3–11, Aldershot: Edward Elgar 1993.
- POPPER, KARL R. (1972), *Objective Knowledge*, Oxford: Clarendon Press.
- QUERVAIN, DOMINIQUE J.-F. DE, URS FISCHBACHER, VALERIE TREYER, MELANIE SCHELLHAMMER, ULRICH SCHNYDER, ALFRED BUCK and ERNST FEHR (2004), "The Neural Basis of Altruistic Punishment", *Science*, 305, pp. 1254–1258.
- RAMSER, HANS-JÜRGEN (1987), *Beschäftigung und Konjunktur*, Berlin: Springer.
- RAWLS, JOHN (1971), *A Theory of Justice*, Cambridge (Mass.): Harvard University Press.
- RIKER, WILLIAM H., and PETER C. ORDESHOOK (1968), "A Theory of the Calculus of Voting", *American Political Science Review*, 62, pp. 25–42.

- RUBINSTEIN, ARIEL (1998), *Modeling Bounded Rationality*, Cambridge (Mass.): MIT Press.
- SAMUELSON, PAUL A., and ROBERT M. SOLOW (1960), “Analytical Aspects of Anti-Inflation Policy”, *American Economic Review, Papers and Proceedings*, 50(2), pp. 177–194.
- SANTAMERO, ANTHONY M., and JOHN J. SEATER (1978), “The Inflation-Unemployment Trade-off: A Critique of the Literature”, *Journal of Economic Literature*, 16, pp. 499–544.
- SCHELLING, THOMAS C. (1978), “Egonomics, or the Art of Self-Management”, *American Economic Review, Papers and Proceedings*, 68(2), pp. 290–294.
- SCHELLING, THOMAS C. (1980), “The Intimate Contest for Self-Command”, *Public Interest*, 60, pp. 94–118.
- SCHELLING, THOMAS C. (1984), “Self-Command in Practice, in Policy, and in a Theory of Rational Choice”, *American Economic Review, Papers and Proceedings*, 74(2), pp. 1–11.
- SCHLICHT, EKKEHART (1977), *Grundlagen der ökonomischen Analyse*, Hamburg: Rowohlt.
- SCHMID, MANFRED (1979), *Handlungsrationalität, Kritik einer dogmatischen Handlungswissenschaft*, München: Wilhelm Fink.
- SCHMID, MANFRED (1979a), „Rationalitätsprinzip und Handlungserklärung“, in: H. LENK (ed.), *Handlungstheorien-interdisziplinär*, Vol. 2(1), pp. 491–533, München: Wilhelm Fink.
- SCHOEMAKER, PAUL J.H. (1982), “The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations”, *Journal of Economic Literature*, 20, pp. 529–563.
- SCHURZ, GERHARD (ed.) (1988), *Erklären und Verstehen in der Wissenschaft*, München: Oldenbourg.
- SCHUMPETER, JOSEF A. (1954), *History of Economic Analysis*, New York: Oxford University Press.
- SEARLE, JOHN R. (2001), “The Classical Model of Rationality and Its Weaknesses”, in: J.R. Searle, *Rationality in Action*, pp. 1–32, Cambridge (Mass.): MIT Press.
- SELTEN, REINHARD (1990), “Bounded Rationality”, *Zeitschrift für die gesamte Staatswissenschaft (JITE)*, 146, pp. 649–658.
- SEN, AMARTYA K. (1977), “Rational Fools: A Critique of the Behavioural Foundations of Economic Theory”, *Philosophy and Public Affairs*, 6, pp. 317–344.
- SEN, AMARTYA K. (1985), “Goals, Commitment, and Identity”, *Journal of Law, Economics, and Organization*, 1, pp. 341–355.

- SEN, AMARTYA K. (1987), “Rational Behaviour”, in: *The New Palgrave, A Dictionary of Economics*, Vol. 4, pp. 68–76, London: Macmillan.
- SEN, AMARTYA K. (1994), “The Formulation of Rational Choice”, *American Economic Review, Papers and Proceedings*, 84(2), pp. 385–390.
- SEN, AMARTYA K. (2005), “Why Exactly is Commitment Important for Rationality”, *Economics and Philosophy*, 21, pp. 5–13.
- SEN, AMARTYA K., and BERNARD WILLIAMS (1982), “Introduction”, in: A.K. Sen, and B. Williams (eds.), *Utilitarianism and Beyond*, pp. 1–21, Cambridge (England) et al.: Cambridge University Press.
- SIMON, HERBERT A. (1955), “A Behavioral Model of Rational Choice”, *Quarterly Journal of Economics*, 69, pp. 99–118.
- SIMON, HERBERT A. (1978), “Rationality as Process and as Product of Thought”, *American Economic Review, Papers and Proceedings*, 68(2), pp. 1–16.
- SIMON, HERBERT A. (1979), “Rational Decision Making in Business Organizations”, *American Economic Review*, 69, pp. 493–513.
- STEGMÜLLER, WOLFGANG (1960), „Das Problem der Kausalität“, in: E. Topitsch (ed.), *Probleme der Wissenschaftstheorie, Festschrift für Viktor Kraft*, pp. 171–190, Wien: Springer.
- STEGMÜLLER, WOLFGANG (1969), *Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, Band I, Wissenschaftliche Erklärung und Begründung*, Berlin: Springer.
- STOCKÉ, VOLKER. (2002), *Framing und Rationalität: Die Bedeutung der Informationsdarstellung für das Entscheidungsverhalten*, München: Oldenbourg.
- SUDGEN, ROBERT (1985), “Why Be Consistent? A Critical Analysis of Consistency Requirements in Choice Theory”, *Economica*, 52, pp. 167–183.
- THALER, RICHARD H. (1992), *The Winner’s Curse, Paradoxes and Anomalies of Economic Life*, New York: The Free Press.
- THALER, RICHARD H., and HAROLD M. SHEFRIN (1981), “An Economic Theory of Self-Control”, *Journal of Political Economy*, 89, pp. 392–406.
- TIETZEL, MANFRED (1981), „Die Rationalitätsannahme in den Wirtschaftswissenschaften, oder: Der homo oeconomicus und seine Verwandten“, *Jahrbuch für Sozialwissenschaft*, 32, pp. 115–138.
- TVERSKY, AMOS (1969), “Intransitivity of Preferences”, *Psychological Review*, 76, pp. 31–48.
- TVERSKY, AMOS, and DANIEL KAHNEMAN (1981), “The Framing of Decisions and the Psychology of Choice”, *Science*, 211, pp. 453–458.
- TVERSKY, AMOS, and DANIEL KAHNEMAN (1987), “Rational Choice and the Framing of Decisions”, in: R.M. Hogarth and M.W. Reder (eds.) (1987),

- Rational Choice: The Contrast Between Economics and Psychology*, pp. 67–94, Chicago: Chicago University Press.
- TVERSKY, AMOS, and RICHARD H. THALER (1990), “Anomalies: Preference Reversals”, *Journal of Economic Perspectives* 4(2), pp. 201–211.
- VANBERG, VICTOR (1975), *Die zwei Soziologien*, Tübingen: Mohr (Siebeck).
- VANBERG, VICTOR (2004), “The Rationality Postulate in Economics: Its Ambiguity, Its Deficiency and Its Evolutionary Alternative”, *Journal of Economic Methodology*, 11, pp. 1–29.
- WATKINS, JOHN W.N. (1970), “Imperfect Rationality”, in: R. Borger and F. Cioffi (eds.), *Explanation in the Behavioural Sciences*, pp. 167–217, 228–230, Cambridge: Cambridge University Press.
- WEBER, MAX (1922), *Wirtschaft und Gesellschaft*, Mohr (Siebeck), Tübingen; English translation: *Economics and Society: An Outline of Interpretative Sociology*, New York: Badminster Press 1968.
- ZINTL, REINHARD (1989), „Der Homo Oeconomicus: Ausnahmeerscheinung in jeder Situation oder Jedermann in Ausnahmesituationen“, *Analyse und Kritik*, 11, pp. 52–69.
- WILSON, EDWARD O. (1999), *Consilience: The Unity of Knowledge*, New York: Vintage.

SUMMARY

The weak rationality principle is not an empirical statement but a heuristic rule for how to proceed in social sciences. It is a necessary ingredient of any ‘understanding’ social science in the Weberian sense. In this paper, first this principle and its role in economic theorizing are discussed. It is also explained why it makes sense to use a micro-foundation and, therefore, to employ the rationality assumption in economic models. Then, we discuss whether the anomalies of individual behaviour as highlighted in modern behavioural economics impair the applicability of the weak rationality principle. This is not the case. We conclude with some remarks on handling the problems of ‘free will’ as well as ‘weakness of the will’ within the economic approach.