Els Opuscles del CREI

num. 45

July 2018

Measures of rationality and welfare for behavioural decision-makers

Jose Apesteguia

CREi

Centre de Recerca
en Economia Internacional

The Centre de Recerca en Economia Internacional (CREI) is a research centre sponsored by the Universitat Pompeu Fabra (UPF) and the Generalitat de Catalunya. It is located at the campus of the UPF in Barcelona.

CREI's objective is to promote research in international economics and macroeconomics subject to the highest standards of academic scholarship.

Els *Opuscles del CREI* are the instruments for the diffusion of CREI's research beyond the academic world. Each *Opuscle*, aimed at a general public, gathers together conclusions and observations of specialized papers published, or about to be published, in professional journals. The opinions stated in *Els Opuscles del CREI* are the sole responsibility of the authors.

Pdf version of this *Opuscle* and all the others *Opuscles del CREI* can be downloaded at: www.crei.cat/opuscles.

Published by: CREI
Universitat Pompeu Fabra
Ramon Trias Fargas, 25-27 - 08005 Barcelona
Tel. 93 542 13 88

Design: Fons Gràfic

ISSN: 2604-2266

Measures of rationality and welfare for behavioural decision-makers[(*)]

Jose Apesteguia[(#)]

1. Introduction

Economics is built on the basis of a model of individual behaviour. Virtually every economic model is grounded on a theory of how individuals make decisions. From the foundations of individual behaviour, economics deals with topics as diverse as household behaviour, organizational behaviour, production, social choice, savings, monetary policy, central banking, trade, financial markets, taxation, education, etc. This individualistic approach represents, in fact, one of the main distinctions between economics and other social sciences like sociology or political science. It comes, therefore, as no surprise that research on individual decision-making has always been a lively field of inquiry in economics. Today, with the interplay of economics, psychology, and the neurosciences, the study of individual behaviour represents one of the most exciting and vibrant areas of economic research.

This *Opuscle* is on the economic model of individual behaviour. In Section 2, I give a short histor-

ical account of the notion of individual economic rationality, and then I lay down the key ingredients of the current, standard view on the subject matter. In a nutshell, the standard view of individual behaviour understands individual decision-making as the outcome of the constrained maximization of a given ranking over all alternatives. It is difficult to conceive of a simpler and more powerful and operational model of behaviour. Undoubtedly, the standard model of individual decision-making has facilitated a tremendous advancement in the understanding of society, and the treatment of difficult economic problems. At the same time, the classical model relies on a number of assumptions about individuals that abstract from important psychological factors that may affect economic decisions in important and systematic ways. As Mullainathan and Thaler (2001) point out, the classical economic model of human behaviour assumes economic agents with unlimited cognitive abilities, unlimited self-control and, in many instances, unlikely selfish behaviour. In Section 3, I review a selection of the behavioural economics literature providing evidence of actual human behaviour deviating from the predictions of the classical model of choice. This is a growing literature that is triggering the development of new models of individual behaviour that incorporate in various ways psychological considerations, and yet are sufficiently tractable to be incorporated in applied economic models.

The evidence reviewed in Section 3 showing that the behaviour of individuals can be systematically at odds with the standard model in economics raises two important questions:

Q.1: How severe are the deviations from the standard model of choice?

Q.2: What is the best way to extract relevant information from the choices of the individual for the purposes of welfare analysis?

These are two fundamental questions and the remainder of this *Opuscle* is devoted to address both of them.

Having a tool to measure the deviations of actual choices from the predictions of the standard model would make it possible to evaluate whether the standard model of individual behaviour is a reasonable way to describe behaviour. That is, the validity of the theory should not be based on whether or not individuals violate the model in a given situation, but on how close their behaviour is with respect to this benchmark. Intuitively, large and systematic deviations would naturally call for the adoption of alternative models of choice. Moreover, the availability of a reliable tool to assess the distance between actual behaviour and behaviour consistent with the standard model will enable interpersonal comparisons. This, in turn, may improve our understanding of individual behaviour and may also prove crucial in the development of future choice models. Furthermore, the possibility of performing meaningful comparisons of rationality will allow evaluation of deviations between various alternative models of choice and, hence, provide a tool to give some structure to the rapidly growing literature on alternative individual decision-making models that are expanding the classical notion of rationality. Section 4 reviews the measures that the literature has offered to evaluate the consistency of individual behaviour with the standard model. There, I also introduce a short review of the main empirical findings obtained by using such measures.

By dealing with Q.2, it would be possible to identify, from an external perspective, the good or bad alternatives for the individual even when the behaviour of the individual is not fully compatible with the classical model of individual decision-making. This is of prime relevance since welfare analysis is at the core of economics. It

would permit, for example, to predict individual behaviour when introducing a change in some economic policy (e.g., a reform in the tax system, changes in health coverage, or a new regulation in the labour market), and evaluate ex-ante what are the policies that may be expected to be optimal from a welfare perspective. In Section 5, I delineate the main contributions in the literature to the question of how to do welfare analysis with individuals that are inconsistent with respect to the standard model.

In Section 6, I present the novel proposal of Miguel A. Ballester and myself on how to measure deviations from rationality and to perform welfare analysis. There, I argue that we offer an integrated approach of the two questions, and explain the advantages of it with respect to the proposals that the literature offers.

Section 7 concludes by suggesting future lines of research.

2. The classical economic model of individual behaviour

The organising principle for the economic view of individual behaviour is the notion of rationality.[1] Rationality is a very difficult concept to delineate. The term rationality means different things to different people, and some argue that it is in fact an unfortunate, narrow label to refer to an economic model of individual behaviour. In the most standard view in economics, rational behaviour means that the individual seeks the best interests that she can attain, given her own view on them. This line of thought belongs to the origins of economic science, and can be found in the writings of Adam Smith, David Hume, and later of Jeremy Bentham and John Stuart Mill. Bentham and Mill advocated the measurement of "happiness" by means of a "felicific calculus", that would be summarised in an utility function, and understood individual behaviour as the result of utility maximization. Bentham and Mill discussed at length the intricate and elusive notion of measuring utility, making interpersonal comparisons of utility, and the like. What is important to remark here is that we have introduced the three key ingredients of the classical model of choice: (1) the individual seeks to maximize (2) her utility function, (3) subject to constraints.

The next significant step in the development of the model came with generation of marginalists, most notably from William Stanley Jevons, Carl Menger and Leon Walras. Jevons, Menger and Walras incorporated the utility maximization approach as a formal theory into a variety of economic models, without paying much attention neither to the measurement of happiness or utility, nor to the establishment of interpersonal comparisons.

Subsequent key developments from Vilfredo Pareto, Irving Fisher, Alfred Marshall, Francis Edgeworth, Eugen Slutsky, and others, laid down the basis for an ordinal view of utility, one that would only care about the ranking of objects, not about the utility intensities attached to the objects. This leads us to the work of Paul Samuelson, who pioneered the next grand step in the evolution of the rational model of choice: the revealed preference approach. It is my view that it is safe to say that the revealed preference approach to individual decision-making represents one of the pillars of modern economics, that has had a profound impact on how economics is done, and, at the same time, it establishes the foundations of economics as an individualistic scientific approach.

The key insight of Samuelson was to note that the results of consumer theory obtained under the

utility maximization assumption can be derived without ever specifying a utility function. Instead of utility, we can treat the choice among a menu of feasible alternatives as a primitive: if we see a consumer choosing an object when there is a second one that he can afford, it must be the case that this consumer prefers the first object over the second. That is, the consumer, by way of her choice behaviour, is revealing a preference for the chosen object over the unchosen, but available ones. Remarkably, Samuelson showed that as long as these revealed preferences display some consistency, the classical results of consumer theory obtained under the utility maximization assumption hold. What are these consistency restrictions on choice? In what follows, I illustrate the type of exercise in the study of consistent, or rational, behaviour in a simple revealed preference framework.[3] Consider the following axiom of choice:

Independence of Irrelevant Alternatives (IIA): Suppose that an individual chooses an alternative x from a set of alternatives B. If S is a subset of B that contains x, then the individual must choose x when confronted with S.

IIA imposes an appealing restriction on individual behaviour, when judged from the perspective of rationality. Namely, if we observe that the decision-maker out of big set B selects alternative x, we can expect that the individual has evaluated all the available alternatives in B, and has come to the conclusion that x is the most preferred alternative in B. If we were now to observe the same decision-maker confronting a smaller set S that only contains some of the alternatives of B, and in particular it contains the chosen alternative from B, alternative x, it makes sense to expect her to choose x, that was shown to be superior to all the alternatives that conform S.

It turns out that, under some technical restrictions, one can show that IIA is all that it is needed to lay down the foundations of the classical individual rationality model. In order to present the result, let me introduce the notion of a preference relation. In the present context a preference relation, or in short a preference, can be understood as a linear ranking over all the possible alternatives, where `linear' here means that there are not two alternatives ranked equally. Preference relations and utility functions are two ways of representing the tastes of an individual. When the only relevant information is the ranking of alternatives, both ways give exactly the same information. We are now in a position to present the following key result in the revealed preference literature:

Theorem 1: Individual choice behaviour satisfies IIA if and only if there exists a preference relation which maximization explains individual choice behaviour.

Let me try to convey the significance of this result. First, it equates the rational model of individual decision-making with the satisfaction of a simple property, IIA, on choice behaviour. This gives testability to the theory, in the Popperian sense. That is, a choice dataset violating IIA cannot be explained by the rationality model, and, on the other hand, whenever a dataset is consistent with IIA, it means that behaviour can be explained as if it were the result of maximizing a preference relation. Moreover, IIA is an easy property to check in practice, facilitating the actual testability of the rational model. Finally, note that the rational model is grounded purely on observed data. There are no assumptions on (unobserved) intensities of preferences, and the like.

I now illustrate the basics of Theorem 1 with a simple example. The example sharply shows what types of behaviour are compatible with the

Table 1. Hypothetical choices

Choice	Available drinks	Customer 1	Customer 2	Customer 3
1	Beer, Wine, Water	Beer	Beer	Beer
2	Beer, Wine	Beer	Beer	Beer
3	Beer, Water	Beer	Water	Water
4	Wine, Water	Wine	Wine	Water

theory, and which are at odds. Suppose that you are the owner of a bar that offers three different drinks: *Beer*, *Wine*, and *Water*. On occasions, due to restrictions on supply, only two of these drinks are available for sale. Table 1 shows the choices made by three different clients, under the different availability situations.

Let us start with the first customer. Note that this customer seems to have a clear preference for *Beer* over all other alternatives. It is chosen whenever it is available, and only when *Beer* is not available, as in choice situation 4, Customer 1 chooses something different, in this case *Wine*. This is exactly the kind of behaviour prescribed by IIA, and hence Theorem 1 implies that the behaviour of this individual can be explained by the maximization of a preference relation. In this case, by the preference relation ranking *Beer* on top of *Wine*, and *Water* and *Wine* on top of *Water*.

Let us now contemplate Customer 2. It is apparent that we cannot reconcile the behaviour of Customer 2 with the model of rationality. Note that while this client is choosing *Beer* when all the possibilities are available, he switches to *Water*, when only *Water* and *Beer* are available. This is a direct violation of IIA, and hence Theorem 1 implies that there is no preference relation which maximization can explain the behaviour of Customer 2.

Finally, Customer 3, although chooses *Beer* over *Wine* and over *Water* in the grand set of alternatives, she always chooses *Water*, whenever this is available in the binary sets of alternatives. Again, this behaviour, although of a different nature to that of Customer 2, cannot be reconciled with the rational model of choice.

I would like to close this brief introduction to the rational model of choice by commenting that economists often attribute to the model both a normative and a positive interpretation. That is, it is taken to represent what an individual should ideally do, and at the same time, it is regarded as a good description of what individuals actually do. For the rational model of choice to be regarded as a good description of what decision-makers actually do, one would wish that there is strong empirical support for the model. This is the question we address in the next section.

3. Behavioural economics

The rational model of individual behaviour introduced in the previous section is based on the maximization principle, in which the alternative chosen by the individual is the one that maximizes a preference relation over the menu of available alternatives. Over the last decades the research has produced increasing amounts of evidence documenting systematic and predictable deviations from this notion of rationality. This literature is the result of the fructiferous interdisciplinary collaboration between economics, psychology and, more recently neurosciences. This body of research constitutes now a sub-field of economics on its own, known as behavioural economics. The insights from behavioural economics are now ubiquitous and permeate other fields of economics, including finance, industrial organization, labour, pub-

lic finance and, more recently, macroeconomics. Behavioural economics aims to develop versatile choice models with more realistic psychological foundations of human behaviour. In this section, I discuss some phenomena that have attracted a great deal of empirical and theoretical attention and prove difficult, if not impossible, to accommodate within the classical model of choice.[4] What follows is a biased selection of behavioural findings. There is, however, no scarcity of very good and readable introductory books; see, e.g., Thaler (2015), Kahneman (2011), or Ariely (2008).[5]

3.1. Cyclical behaviour

Acyclicity is a crucial assumption underlying the classic rational choice. Assume an individual that chooses alternative x over alternative y and y over alternative z. Acyclicity implies that it cannot be the case that z is chosen over x. Otherwise note that from the menu composed of alternatives x, y and z, there is no best alternative since every alternative is dominated by another alternative, and hence there is no obvious candidate for choice, posing a serious challenge to the classical rational choice model. Clearly, whatever the choice from x, y and z an IIA violation will follow. Acyclicity forbids this kind of binary cycles involving any number of alternatives.

Experimental evidence shows that acyclicity is frequently violated in individual choice. Already in 1954 May conducted a class experiment where he asked his students to choose hypothetical marriage partners that differed in three dimensions: wealth, beauty and intelligence. Partner 1 was the wealthiest, had intermediate beauty, and showed the lowest ranked in intelligence. Partner 2 was the most beautiful, intermediate intelligence, and the lowest wealth. Finally, partner 3 was the most intelligent, had intermediate wealth, and was the last in terms of beauty. He documented that the preferences of about a third of his students resulted in cycles. That is, they would exhibit behavioural patterns of the sort of preferring Partner 1 over 2, 2 over 3, and yet 3 over 1. The key lesson to note here is that when the comparison of alternatives involves taking into consideration different dimensions and there are no clear dominations, the choice task may be a complex one, which facilitates the violation of acyclicity. Other papers showing violations of acyclicity involving multi-dimensional objects are Tversky (1969), using lotteries (with dimensions involving probabilities and prizes), and Roelofsma and Read (2000) using inter temporal choices (monetary prizes and time).

Another related phenomenon that has attracted a good deal of attention is the so-called preference reversals (see, among many papers, the works of Lichtenstein and Slovic (1971); Gret and Plott (1979); Tversky (1990); Loomes, Starmer and Sugden (1991)). A typical experiment involves pairs of lotteries with probabilities over two monetary outcomes, a positive one and zero. In each of these pairs, one lottery (the P-bet) offers a relatively large chance of a modest prize, while the other (the $-bet) offers a smaller chance of a larger prize. In the first stage, subjects are asked to choose between the P-bet and the $-bet. In a second stage, a monetary valuation for each of the lotteries is elicited. It has been systematically shown that people often choose the P-bet in the first stage while attach a higher monetary value on the $-bet. Obviously, this behaviour is very difficult to reconcile with the model outlined in the previous section.

3.2. Framing

The framing of a decision problem encompasses the details of the description or presentation of the objects of choice, and all the features around the decision process. Often one can change the

framing of the decision problem in an inconsequential way, in the sense that the alternatives are identical under the different frames. For the classical rational choice model, these changes in the frames should not affect behaviour, since at the end of the day the alternatives are exactly the same, and so are the consequences of choice. Yet, there is ample evidence establishing framing effects, that is the dependence of individual behaviour to the particular framing of the decision problem.[6]

I illustrate the framing effects with a classical experiment by Tversky and Kahneman (1981). A group of individuals was presented with the following hypothetical decision problem: "Imagine that the U.S.A. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programmes to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programmes is as follows:

- If Programme A is adopted, 200 people will be saved.

- If Programme B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved."

A second group of individuals was presented with exactly the same cover story of the previous problem, but with a different presentation of the alternative programmes:

- "If Programme C is adopted, 400 people will die.

- If Programme D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die."

It is immediately apparent that programmes A and C are identical in terms of the consequences, and so are programmes B and D. They only differ in the details of the presentations. Accordingly, we should expect a comparable number of people choosing A over B in the first presentation of the problem, to that choosing C over D in the second presentation. Yet, a large majority of participants chooses programme A over B (72% versus 28%), and a large majority selects programme D over C (78% versus 22%). This is an astonishing result that has been replicated many times, with very different subject pools, involving students, managers, medical doctors, policy-makers, etc. It shows that a subtle change in the wording of the decision problem may have a large impact on behaviour. In this particular case, the first presentation emphasises the gains dimension, that is the saving of people, while the second presentation emphasises the losses dimension, the dying of people. There is ample evidence that human beings behave very differently when confronting gains versus losses; while gains trigger risk aversion, losses are associated with risk loving behaviour. This behavioural pattern is consistent with the above results; in the first presentation a majority of people prefer the safe option A, while in the second presentation, people seem to prefer the risky lottery over the safe option, probably motivated by the possibility of avoiding all deaths.[7] There are many other instances of framing in the literature, involving a diverse array of choice situations and subject pools. The interested reader can consult Tversky and Kahneman (1981) and Thaler (1999).

3.3. Menu effects

Property IIA, discussed in the Introduction, implies that if one chooses x over y, when only x and y are available, one should not switch to y when a third alternative z is added to the menu. This would imply a direct violation of the property, and

hence, of the classic rational choice. However, research has identified a variety of violations of IIA of this sort that have been named as menu effects. Here I shall present three such cases.

In the attraction effect (Huber, Payne and Puto (1982); Simonson (1989)) the alternatives have different values on two desirable attributes (e.g., speed and safety of a car). Participants in the experiment first confront two alternatives, x and y, where x dominates y in terms of one of the attributes and is dominated by y with respect to the other attribute. Hence, no option dominates the other on both dimensions. Some individuals choose x over y and some the reverse. Then, the experimentalists add a third alternative z that is dominated by y in the two attributes, but dominates x in one attribute and is dominated by x by the other attribute. The results are that the introduction of z in the menu makes a significant fraction of people to switch to option y. It seems that the fact that y dominates z, but x does not, makes y to appear more attractive.

The compromise effect (Simonson 1989) shows that people have the tendency to select intermediate options. As an illustration, taking the x and y options above and adding a third alternative h that is in between x and y in the two attributes, makes individuals to turn to the choice of h. Interestingly, if to the menu composed of x and y another third alternative g is added now makes y fall in between x and g in the two attributes, the tendency is that people turn to the choice of y.

A third menu effect to which I want to refer here is the so called choice overload effect (Iyengar and Lepper 2000). In a typical experiment, a group of individuals is offered a small number of options (e.g., 6 different flavours of jam) and another group is offered a large number (e.g., 24 different flavours). The results are that individuals are more likely to choose an option (to buy a jam) in the first case, than in the second. The standard interpretation is that too many options in the menu may be demotivating, making individuals to abstain from making a decision.

These effects show that choice behaviour can be malleable by manipulating the composition of the menu of options in specific ways. Clearly, these findings are at odds with the predictions of the rational model of choice.

3.4. Inattention

The rational economic model of human behaviour assumes economic agents with unlimited cognitive abilities. This is obviously an idealization of human behaviour. It comes as no surprise that research has documented that the attention we pay to certain objects, or attributes of objects, can be easily manipulated, with behavioural consequences that are at odds with the predictions of the standard model.[8]

There is overwhelming evidence that people tend to pay more attention to the objects that are presented first in a list. For example, Itzkowitz, Itzkowitz and Rothbort (2016) show that early alphabet stocks are traded more frequently than later alphabet stocks. Einav and Yariv (2006) find that in economics, where the order of authors in the academic papers typically respect alphabetical order, authors with earlier surname initials do better in terms of promotions and competitive prizes.

Another effect that is attributed to inattention is the left-digit bias, by which individuals focus on the number's leftmost digits. There is plenty of anecdotal evidence for this in the pricing strategies of firms. An interesting illustration of the bias is due to Lacetera, Pope and Sydnor (2012) that study the second-hand car market. They show prices, by the

tendency of buyers on focusing on the left digits on the odometer, present discontinuous falls at 10,000-mile thresholds, along with smaller drops at 1,000-mile thresholds. This shows that an increase in a few miles keeping the left digit in the odometer unchanged has a much smaller effect on prices than when the increase in miles makes the first digit in the odometer to increase.

In the U.S.A., product tags at the supermarket show pre-tax prices, and local sales taxes are added at the counter. In a fully attentive scenario, this procedure would be inconsequential since consumers anticipate the final price to be paid. Chetty, Looney and Kroft (2009) show that this is not the case. When they included the final price in the product tags, consumption dropped by about 8%.

3.5. Reference dependence

Human beings often take reference points in order to judge the different alternatives.[9] This is manifested in a variety of settings. In the status quo bias, the alternative that is set as default is typically valued more highly than the others. Kahneman, Knetsch and Thaler (1990) show that the monetary evaluation of an object (in the experiment, a coffee mug with the logo of the university) depends crucially on whether the object has been given to the individual or not. In the first case, the object is valued almost double than in the second case.

Reference dependent behaviour can have very important welfare implications. Madrian and Shea (2001) study the impact of automatic enrolment on 401(k) savings behaviour, one of the most important sources for retirement savings in the U.S.A. In a typical 401(k) plan, a worker determines her retirement savings contributions, and often the employer proportionately matches employee's contributions. The authors find that 401(k) participation increased significantly under automatic enrolment, as opposed to the case where workers have to make an active choice of the 401(k) they want to get involved in. The experiment consisted in letting a group of workers choose their preferred plan, and, as usual, if there is no active choice of a plan, the worker does not enrol in any. Another group of workers were automatically enrolled in one plan, and had the option of opting out whenever they wanted. As mentioned, automatic enrolment markedly increased participation in 401(k) plans, which may have large welfare consequences at the retirement age.

Further cases of reference-dependent behaviour can be found in other economic and social settings. The housing market is one interesting case. Using data from downtown Boston in the 90s, Genesove and Mayer (2001) find that the sellers whose selling price falls below their original purchase price (in nominal terms) set an asking price that exceeds the asking price of other sellers by between 25 and 35 percent of the percentage difference between the two. Moreover, they exhibit a much lower sale hazard than other sellers. This is clearly another instance where reference dependent behaviour may have large economic consequences.

4. Measures of rationality

The preceding section suggests that the evidence against the classical model of rationality is overwhelming. As a matter of fact, the type of behavioural empirical findings reviewed above have triggered the development of a wealth of alternative theoretical models of decision-making, incorporating in various ways some of the behavioural phenomena.[10] However, for our purposes in this *Opuscle*, it is crucial to note the difference between

establishing inconsistencies with the classical rational model and evaluating the extent of these inconsistencies. The idea here is to discriminate between violations that fall somehow 'close' to the model, from those that signify a "sizeable" departure. Of course, to make such judgements, we need a proper tool to measure the distance between actual behaviour, and the behaviour predicted by the rational model. I devote this section to review the different ways that the theoretical literature has suggested to deal with this challenge, and the empirical findings that we have gained with them.

The main theoretical proposals in the literature are due to Afriat (1973), Houtman and Maks (1985), and Swofford and Whitney (1986). These papers suggest measurement techniques that approach the problem from very different angles. Let me now explain the basic ingredients of these measures, starting with the first and most prominent measure of inconsistency, the Afriat's index, after Afriat (1973). In the setting of Afriat, the menus of alternatives are defined by the prices of the goods and by the wealth of the consumer. These two components determine the goods the consumer can buy. In the presence of inconsistent behaviours across menus, the suggestion of Afriat is to compute the percentage of wealth reduction that would make all the inconsistencies to disappear. The idea is that by hypothetically reducing wealth, the menus of available alternatives shrink (i.e., there are less goods that the consumer can afford), and with them some of the inconsistencies disappear. Then, Afriat's measure of inconsistency corresponds to the minimum proportion of wealth reduction such that all the inconsistencies in the data vanish. One of the main advantages of Afriat's index is that it is easy to compute, even in large datasets. The main disadvantage of the index is that by focusing on the maximum inconsistency present in the data it is not sensitive to the number of inconsistencies.[11]

A second influential proposal is due to Houtman and Maks (1985). The suggestion here is to measure the inconsistency of a dataset by the minimum number of violations that are needed to be discarded for the remaining dataset to be consistent with the classical model of rationality. The positive aspect of this proposal is that it responds to the number of inconsistencies present in the data. The disadvantage is that it does not evaluate the severity of each inconsistency; all inconsistencies count the same in the final judgement of rationality, independent of their severity.[12]

Yet, a third approach, due to Swofford and Whitney (1986), entails counting the number of violations of a rationality axiom, say IIA. This is another natural way of measuring the rationality of individuals. The disadvantage is that, like in the former case, by counting the number of violations of an axiom, its judgement is not necessarily aligned with the severity of the inconsistencies.[13]

We now turn to the empirical findings obtained in the measurement of rationality. In order to facilitate comparisons between the different findings, here we will focus on two particular measures: (i) the average Afriat's index, and (ii) the percentage of individuals in the different studies whose behaviour cannot be explained by the rational model. Table 2 summarizes the results from a variety of recent studies, involving different subject-pools, different types of goods, different data sources, and different experimental methodologies.[14]

The papers cited in Table 2 follow a revealed preference approach, in the sense of studying the case of a group of individuals, where every individual selects her preferred options from different menus of options. The type of individuals, options and sets of options differ from study to study. Choi, Kariv, Müller and Silverman (2014) perform a large-scale experiment involving 1182

Table 2. Summary of the empirical literature on the measurement of rationality

Study	Sample	Choice Context	Average Afriat's Index	Percent of Inconsistent Individuals
Choi et al. (2014)	Representative sample of 1182 individuals from Netherlands	Choice under risk	0.12	77%
Harbaugh et al (2001)	31 second-grade students, 42 sixth graders and 55 college undergrads	Consumption bundles (chips and juices)	0.07	74% of second graders 38% of sixth graders 35% of undergrads
Andreoni and Miller (2002)	176 students of upper-level economic courses	Altruism	0.01	10%
Fisman et al.(2015)	208 students of Yale Law School (YLS) and 309 young adults from ALP	Altruism	0.05	≥25% of YLS ≥30% from ALP
Carvalho et al. (2016)	1119 U.S. low-income households	Choice under risk	0.15	80%
Echenique et al. (2011)	494 households from U.S. mid-western city	Consumption bundles (grocery store)	0.02	81%
Dean and Martin (2015)	977 households in Denver metropolitan area	Consumption bundles (grocery store)	0.01	71%

individuals, representative of the Dutch population. In the experiment, each individual is presented with a sequence of 25 decision problems where at each decision problem a lottery must be selected from a menu of lotteries. Harbaugh, Krause and Berry (2001) study rational choice in 31 second-grade students (approximately 7 years old), 42 sixth graders (11 years old) and 55 college students (21 years old). Participants in this study are presented sequentially with 11 choice sets. Each choice set consists in different bundles of small bags of potato chips and of boxes of fruit juice. Participants choose one bundle from each one of the offered sets of bundles. Andreoni and Miller (2002) and Fisman, Jakiela, Kariv and Markovits (2015) study the consistency of altruistic behaviour. In both studies, participants are presented with a sequence of decision problems where subjects have to distribute a monetary endowment between themselves and another anonymous participant, where the possibilities of distribution vary from one decision problem to another. In Andreoni and Miller (2002), the test is carried on a pool of 176 volunteers of upper-level economic course, while the subject pool of Fisman, Jakiela, Kariv and Markovits (2015) is composed by 208 students of Yale Law School. Carvalho, Meier and Wang (2016) study the consistency of behaviour of 1119 low-income U.S. households with respect to decisions involving risk and time. Participants choose a lottery from each one of 25 different sets of lotteries and an intertemporal money allocation from 12 different possible allocations. Echenique, Lee and Shum 2011 study scanner data of food expenditures in a panel of 494 households in the U.S.A. Finally, Dean and Martin (2015) study a panel of supermarket data for 977 households in Denver metropolitan area from 1993 to (1995).

The picture that arises from the results of these studies is that the percentage of people being inconsistent can be very large, but, importantly, it seems that the average extent of their violations seem to be relatively small. This is for subjects as diverse as seven year-old children to Yale law students, and studies that employ laboratory experiments, large-scale experiments or panel data of supermarket purchases. Note that in Choi, Kariv,

Müller and Silverman (2014), Echenique, Lee and Shum (2011) and Dean and Martin (2015), the percentage of inconsistent subjects is notable, above 70%. At the same time, it is the case that in all the reviewed cases, the Afriat's index is relatively close to 0, the case of perfect rationality, suggesting that the extent of irrationality may not be large. Note, however, the limitation of Afriat's index commented above. Afriat's index, by construction, only focuses on the maximum inconsistency of the data and is not responsive to the number of inconsistencies. This implies that what these studies are really showing is that the average individual maximum inconsistency is relatively minor. An interesting exercise remains to be done; the analysis of all these datasets using other measures than Afriat's.

In the previous paragraph, we have focused on average effects. However, it is probably the case that the most interesting questions regarding the understanding of inconsistent behaviour arise when accounting for heterogeneity. What explains the different degrees of rationality across individuals? While there is no definite answer to this question, current studies shed some light on potential variables associated with rationality. The consistency of choice seems correlated with socioeconomic variables like income and education, and demographic variables like age. There is some consensus in that high-income and high-education subjects display greater consistency than lower-income and lower-education subjects. Moreover, young subjects tend more toward utility maximization than older ones. I believe that there is still a lot to learn from this kind of analysis. I will return to this point in the concluding section.

5. Behavioural welfare analysis

Welfare economics represents an important area in economics. It is concerned with the evaluation of the welfare of a group of individuals, say a society. The approach entails advocating a particular way of judging different options from a societal perspective that depends crucially on the preferences of the individuals in the society. As an illustration of a typical exercise in welfare economics, consider the evaluation of different distributions of resources in society that may arise, for example, as the result of different tax systems. Individuals in society have different preferences over the distributions of resources, and the question arises on how to actually distribute them. The literature offers many approaches to this problem. A cornerstone one is to focus on the Pareto efficient outcomes. These are the distributions of resources in which there is no individual that can improve in her preference ranking, unless some other individual is made worse off.

What is clear from the above discussion is that at the heart of welfare economics is the assumption that there is a preference relation accounting for individual behaviour. The challenge arises when individuals are inconsistent, and hence their behaviour cannot be explained by way of the maximization of a preference relation. If, for example, the individual on occasions chooses an option x over another y, and in other occasions chooses y over x, the ranking of x and y from an external perspective is not immediate anymore, and hence, the standard tools to perform welfare analysis, both at the individual and societal level, are not well suited. Behavioural welfare analysis is the branch in economics that studies how to establish welfare judgments when individuals are inconsistent with the rational choice model. The purpose is to come with an individual welfare

ranking that ultimately could serve as the basis for both, individual and societal welfare analysis.

There are two main approaches to the question of how to extract the relevant information for welfare considerations in these situations. The first approach involves a choice model-free view, while the second builds on a particular model of decision-making consistent with the behaviour of the individual. In the choice model-free approach, the analyst is agnostic about the underlying choice model the individual may follow and focuses exclusively on the observable choice data. That is, no particular assumption on the sources of the choice inconsistency is made, and revealed choices are taken as the unique source of information to infer individual welfare. The main motivation for this approach is that the analyst may feel that she does not have enough information to be sure about the particular model the individual is adopting. Accordingly, the analyst prefers a methodology that is general enough to be applicable to whichever rationale is used in the process of decision-making. This is the main perspective adopted by the pioneering work of Bernheim and Rangel (2009).[15] Alternatively, the analyst may assume a particular boundedly rational choice model, consistent with the revealed choices of the individual, and use this model in order to infer individual welfare. For example, one could assume that the individual behaves according to the reference-dependent model of prospect theory, and use this particular model to infer the closest welfare ranking. The underlying assumption is that the causes of the inconsistencies may provide valuable information about how to fix the welfare ranking. For example, in the prospect theory case, knowing the reference-dependent nature of behaviour may be informative in order to build an appropriate welfare ranking. This is the approach taken by Green and Hojman (2007).[16]

Interestingly, Bernheim and Rangel (2009) and Green and Hojman (2007), following the above different methodological approaches, independently suggest the use of the same welfare notion. Let us denote by P^B the Bernheim-Rangel-Green-Hojman welfare preference, defined as xP^By whenever there is no observation in the dataset where y is chosen from a menu of alternatives containing alternative x. That is, xP^By whenever there is no piece of evidence in the choices of the individual placing y above x. Bernheim and Rangel (2009) show that, under certain assumptions, P^B is acyclic, and hence it is consistent with the maximization principle, in the sense that P^B will always identified one, or several, best alternatives. This represents a conservative welfare criterion in the sense that it focuses on those pairs of alternatives for which there is no contradictory evidence, and leaves the case of problematic pairs of alternatives as unranked.

We now turn to the empirical studies motivated by these theoretical proposals. Bernheim, Fradkin and Popov (2015) apply the fundamentals of Bernheim and Rangel (2009) to study the behavioural welfare implications of the default options in 401(k) plans. Recall the power of defaults, as reviewed in Section 3. There, I have argued that it has been shown that setting a 401(k) plan as default markedly increases contributions, as compared to a setting without defaults and with the need of an active choice for one plan. Bernheim, Fradkin and Popov aim to study the welfare consequences of these interventions. Interestingly, they show that the optimal default option from the perspective of behavioural welfare is at the extremes; either at the highest contribution rate matched by the employer, or at zero. Bernheim, Fradkin and Popov also study the behavioural consequences of a number of policy interventions, like establishing penalties for not making a choice, and quantify the welfare stakes under different versions of the defaults.

Bouacida and Martin (2015) provide the first direct empirical test of the Bernheim-Rangel-Green-Hojman welfare preference P^B. In one of the tests, Bouacida and Martin use a consumer setting, where P^B does not need to be acyclic, and hence may be uninformative for the purpose of selection of maximal alternatives. Bouacida and Martin show that although all households are inconsistent, the welfare criterion P^B is acyclic in the vast majority of cases (80% of households), and hence useful in the identification of best alternatives.[17]

6. The swaps index

In the previous two sections, we have reviewed the literatures involving the measurement of rationality and the analysis of behavioural welfare. It is noteworthy the separation of the two literatures, despite the fact that the two questions are intimately linked; after all, it is the presence of inconsistent behaviour with the rational model what generates problems in the study of individual welfare. In this section, I review the approach I have developed together with Miguel A. Ballester, Apesteguia and Ballester (2015), in order to jointly address the measurement of rationality and welfare of individuals whose behaviour is inconsistent with the standard model of rationality. In a nutshell, we offer an index that it is sensible to the number and severity of the inconsistencies, and that it also provides a tool for welfare analysis. The severity of the inconsistency is measured by the number of good alternatives the individual has missed in her choice. In what follows, I explain in more detail our approach, while keeping the technicalities low.

Consider a given preference relation P, that is a ranking over the set of alternatives, and one observation (A,a) in the behavioural dataset, representing the choice of alternative a from the menu of available alternatives A. If alternative a is the maximal alternative in menu A according to the preference relation P, then P rationalizes the choice, and there is no inconsistency. The question arises on how to ponder the case when a is not the maximal alternative in A from the perspective of P. We would like to introduce a way of discriminating cases where a is close to be maximal, representing a relatively small inconsistency, from cases where a is far from maximality. Furthermore, we would like to do so without making particular parametric assumptions on preferences. The novel way we propose is to focus on the number of alternatives in the menu A that are preferred to alternative a according to the preference P. If there are any, these are the alternatives that need to be swapped with the chosen alternative a in order for P to explain the choice of a from A. Note the intuition behind this exercise. The number of alternatives that need to be swapped represents the extent of the inconsistency of preference P with the observation (A,a). This has the advantage of evaluating the importance of the inconsistency by discriminating between cases where a is high in the ranking P, from those where a is low, which has a direct welfare interpretation.

In our swaps index, we adopt the above methodology and evaluate all the observations in the behavioural dataset in order to find the preference relation that minimizes the required number of swaps. That is, different preference relations will require different swaps, and we suggest to focus on the preference relation that requires the least number of swaps in order to rationalize choice behaviour. The identified preference relation is the one that best explains the choices of the individual, and the number of swaps that the identified preference relation requires to rationalize the data represents the actual measure of rationality.

The advantages of our approach are the following:

- With respect to the measurement of rationality, the swaps index is sensible to both, the number of inconsistencies and their severity. Let me illustrate with a couple of simple examples. The examples help to appreciate the type of exercise involved in the assessment of rationality judgements involved in the swaps index, and other key indices.

Example. Consider the set of alternatives $X=\{x,y,z\}$. Suppose there are two individuals, 1 and 2, that confront relatively often the binary sets, say 10 times each binary set. Suppose further that the behaviour of the two individuals in the binary menus is totally consistent with the maximization of the preference relation P that regards x as better than y, and y as better than z. Suppose that, in addition, each individual confronts once the set X, and here their choices diverge. While individual 1 chooses option y, individual 2 selects alternative z. Clearly, the choices of both individuals are inconsistent with the preference P, that rationalizes the big bulk of the respective datasets. At the light of preference relation P, it seems logical to argue that the inconsistent choice of individual 1, namely the choice of y from menu $\{x,y,z\}$, is less severe than that of individual 2, that chooses z from $\{x,y,z\}$, since the latter selects the worst possible alternative in P, while the former selects the intermediate one. This is exactly the judgement of the swaps index. Note, however, that other indices do not accommodate these considerations. For example, according to Houtman-Maks' index, eliminating the observation related to menu $\{x,y,z\}$ from the behavioural datasets of the two individuals makes the remaining observations compatible with the model of rationality, and hence the inconsistency index associated to both individuals is equal to 1. Hence, Houtman-Maks's index is unable to discriminate between the inconsistency revealed by the two individuals.

Example. Suppose again two individuals, 1 and 2, repeatedly choosing from various compositions of the menu of alternatives $X=\{x,y,z\}$. Let us assume that both individuals have made the following choices: they have chosen (i) 10 times x from X, (ii) 10 times x from $\{x,z\}$, (iii) 10 times y from $\{y,z\}$ and (iv) 100 times x from $\{x,y\}$. Clearly, these choices are consistent with the maximization of the same preference than in the previous example, $xPyPz$. In addition, suppose that we also observe one extra choice of each individual: individual 1 chooses y from X and individual 2 z from X. Exactly the same arguments used in the previous example show that the two individuals are inconsistent, and that, at the light of preference relation P that rationalizes a big portion of the data, individual 2 is more inconsistent by committing a larger mistake in menu X by choosing z, instead of the optimal alternative x. Let us consider now the index that counts the number of IIA violations (the IIA-index à la Swofford-Whitney, as introduced in Section 4). With respect to individual 1, note that the 10 choices of x from X together with the choice of y from X generate 10 violations of IIA. Similarly, another pair of observations involved in IIA-violations are $(\{x,y\},x)$ and (X,y), and since there are 100 observations of the first type and 1 observation of the second, these observations generate 100 IIA violations. Both types of violations give a total of 110 IIA violations for individual 1. With respect to individual 2, it is now easy to see that the data involved in IIA violations are the 10 observations of x from X with the observation of z from X, the 10 observations of x from $\{x,z\}$ with the observation of z from X, and the 10 observations of y from $\{y,z\}$ with the observation of z from X, leading to a total of 30 IIA violations. Hence, according to the IIA-index, it is individual 1 who is more inconsistent. This, as we have ar-

gued, is in sharp contrast with the conclusion of the swaps index.

- With respect to welfare analysis, the swaps index internalizes the whole dataset in order to find the closest preference relation to the choices of the individual, while other approaches only consider part of the data. Let me illustrate the difference in the treatment of welfare of the swaps index with respect to the approach of Bernheim-Rangel-Green-Hojman introduced in Section 5. In order to do so, we use the following stylized example:

Example. Consider a dataset composed by the following observations: (i) 10 choices of x from $\{x,y\}$, (ii) 10 choices of y from $\{y,z\}$, (iii) 1 choice of y from $\{x,y,z\}$, and finally (iv) 1 choice of z from $\{x,z\}$. Clearly, in the Bernheim-Rangel-Green-Hojman approach, it is zP^Bx, since x is never chosen in the presence of z. However, to evaluate the ranking of alternatives x and z, the swaps index considers the whole dataset. The 10 observations of x from $\{x,y\}$ together with the 10 observations of y from $\{y,z\}$ signify a strong argument for the preference $xPyPz$. This preference implies a mistake in the choices of y from $\{x,y,z\}$ and of z from $\{x,z\}$, but rationalizes the more frequent evidence of the choices of x and y from $\{x,y\}$ and $\{y,z\}$, respectively. Preference P is in fact the optimal preference relation for the swaps index P_S. Hence, we observe that the Bernheim-Rangel-Green-Hojman welfare ranking, by considering a subset of the dataset, may rank alternatives in the opposite direction than the swaps index.

- An important feature of the approach here suggested is that, for the first time, it integrates in the same framework the measurement of rationality and welfare. The analysis of the swaps index identifies a preference relation as the closest to the choices of the individual, and a distance of this preference relation with respect to the mentioned choices. Intuitively, a relatively small distance enhances the confidence on the identified preference relation as a sound basis for welfare analysis. Alternatively, a large distance represents a call of caution regarding the proper identification of the welfare ranking of the individual, and in favour of the consideration of other methods. A natural approach in the latter case is to attempt to learn whether there are stable and predictable regularities in the choice inconsistencies that could be incorporated in the welfare analysis. This is, roughly speaking, the approach advocated by the papers cited in footnote 16.

- Another advantage of the swaps index is that in Apesteguia and Ballester (2015) we provide the axiomatic foundations of the index. That is, we identify a set of properties on inconsistency indices that uniquely identify the swaps index. This represents the first axiomatic exercise in this literature, and sets the swaps index in a solid theoretical framework. Moreover, we also provide the axiomatic foundations for other measures of rationality, like those of Houtman and Maks (1985) and Varian (1990). Our analysis permits the understanding of the properties underlying the measurement of rationality and welfare, and permits the further theoretical development of indices, by investigating alternative properties to the ones we identify in our paper.

One may wonder about the possible differences in the understanding of rationality that the different indices bring. In order to address this question, we take the data of Harbaugh, Krause and Berry (2001), described in Section 4. We first note that, using the 58 participants whose behaviour is inconsistent with the rational model, the correlation between the swaps index and Afriat's index is rather low (.51), while with Houtman-Maks and with the IIA-index is rather high (.97 and .83, respectively). It emerges, therefore, that the conclu-

sions reached with the swaps index and the most used rationality index in the empirical literature, Afriat's index, may differ substantially. In addition, by inspecting closely each individual case, we identify a number of sizable discrepancies in the rationality judgement between the swaps index on the one hand and Houtman-Maks and the IIA-index on the other. In order to understand the source of discrepancies, let me focus on the index of Houtman-Maks. There are subjects in the experiment with a relatively large number of inconsistencies that are regarded as very inconsistent by the index of Houtman-Maks, but where the inconsistencies are relatively minor. At the same time, the experimental dataset contains a number of participants with fewer number of inconsistencies, but arguably, the inconsistencies of these individuals are severe. These later individuals are regarded by Houtman-Maks' index as less inconsistent than the former ones. The swaps index, however, by integrating both types of considerations, namely the number and the severity of the inconsistencies, provide, what we belief is, a more appropriate treatment of these subjects, which implies reversing their inconsistency rankings. Note that an analogous argument can be made with respect to the number of IIA violations.

Let me close this section by commenting on a broad generalization of the swaps index that we undertake in Apesteguia and Ballester (2015). The basic idea of the swaps index is to consider every possible inconsistency between an observation (A,a) and a preference relation P through the number of alternatives in A that are above a according to P. However, instead of considering just the number of alternatives that need to be swapped, we could condition our rationality and welfare judgment on the nature of the menus of alternatives or/and the nature of the chosen alternatives. For example, we may consider important to condition on the size of the menu, or the cardinal utility values of the alternatives that need to be swapped.

7. Final remarks

In the same way, we measure weight and temperature we may aspire to measure rationality. Here we have discussed the approaches that the literature has suggested to measure the consistency of individual behaviour with respect to the rational model of choice. This facilitates a deeper understanding of individual behaviour, giving rise to new questions.

From an empirical perspective, apart from the questions already commented in Section 4, a sound distance measure of rationality would permit to address inquiries such as (i) which individuals are more rational, (ii) when and where individuals are more rational, (iii) what external and internal factors influence rationality, (iv) whether rationality can be learnt, (v) the influence of illness on rationality, etc. These are critical questions that the literature has so far neglected, but that seem essential for a proper understanding of human decision-making. I am convinced that we will see growing work on the subject matter in the near future. This, seems to me, represents a fascinating area for future research, one that lies at the intersection of theoretical and empirical economics, psychology, neurosciences and sociology.

From a theoretical perspective, rationality indices allow to introduce a metric to compare the different behavioural and bounded rationality models of individual behaviour with respect to the canonical model of rationality. This is important, as it introduces discipline in the comparison of theoretical models of individual behaviour. Also, future work may extend the rationality indices to consider

other reference models than the classical model of choice. For example, we may be interested in how far is actual behaviour with respect to prospect theory or with respect to hyperbolic discounting. Also, a challenge that most of the inconsistency indices discussed in this paper face is the difficulty of applying them in settings involving many alternatives. In the language of computer science, most of the rationality indices are NP-complete; that is, they become impractical as the number of alternatives grows. Future research should explore the connection with the algorithmic computer science literature, for the development or adaptation of algorithms that give reasonable approximations to the measurement of rationality. Some steps in this direction have been undertaken by Alcantud, Matos and Palmero (2010), Smeulders, Spieksma, Cherchye, and De Rock (2014), and Apesteguia and Ballester (2010), (2015).

We have also discussed several procedures for identifying welfare rankings, when the individual is inconsistent. The challenge from this perspective has proved to be in the transition from the theoretical world to the applied one, since there is a scarcity of empirical works. This is unsatisfactory, since the ultimate purpose of the theoretical exercises studied in this *Opuscle* is to offer a tool to the policy-maker or the social planner to be used when making decisions on behalf of possibly inconsistent decision-makers. Clearly, this represents a challenge of high scientific interest, since it may help to take better informed decisions when attempting to increase social welfare.

Notes

(*) This Opuscle is based on my joint paper with Miguel A. Ballester, "A Measure of Rationality and Welfare". I am deeply indebted to Miguel for our collaboration on this project in particular, and on all the projects we have undertaken together over the last fifteen years. I would also like to thank Angelo Gutierrez for superb research assistance.

(#) ICREA, Universitat Pompeu Fabra and Barcelona GSE. E-mail: jose.apesteguia@upf.edu.

(1) In preparing parts of this section I have used the following sources, that I highly recommend to the interested reader: Samuelson (1938), Mas-Colell (1982), Mas-Colell, Whinston and Green (1995), Blume and Easley (2008), Nasar (2011), and Chambers and Echenique (2016).

(2) Of course, both objects could be equally desirable for the consumer. In this document, for the sake of exposition, I will abstain from this possibility, and assume the individual selects one single object from a given menu of objects, and that for every pair of objects, the individual always strictly prefers one over another. The reader should know, though, that the issue of indifferences is well incorporated in the formal apparatus.

(3) I assume a finite number of alternatives, and, as commented in the previous footnote, single-valued choice functions. The interested reader can find a discussion of the required technical assumptions and proofs of different versions of Theorem 1 below in Mas-Colell, Whinston and Green (1995) or in Rubinstein (2012).

(4) Naturally, the first reactions of the profession to the findings of behavioural economics were of scepticism. Logically plausible arguments were used that would make the claimed inconsistent behaviour, consistent with the rational model. Of course, this belongs to the normal scientific debate, expected when the current paradigm is challenged. Elaborating further on this debate, albeit important and interesting, would make us deviate from the main objective of this Opuscle. *For the interested reader, Thaler (2015) gives an excellent account of these arguments, and the counterarguments used by behavioural economists.*

(5) Among the important behavioural literatures that I omit, here are the ones referring to social, risk, and time preferences.

(6) To see the connection of framing effects with IIA, suppose alternative x *is chosen from menu* A *and alternative* y *is different to* x *is chosen from menu* A', *where* A' *is the same menu than A but framed differently. This behaviour represents a direct violation of IIA.*

(7) This gains-losses behavioural dichotomy is at the heart of prospect theory of Kahneman and Tversky (1979), which is regarded as the most successful model in the behavioural literature.

(8) The connection drawn between framing effects and IIA violations in Section 3.2 applies here with minor modifications.

(9) Again, the connection drawn between framing effects and IIA violations in Section 3.2 applies here with minor modifications.

(10) Some of the key works are Kahneman and Tversky (1979), Rabin (1993), Laibson (1997), Fehr and Schmidt (1999), Masatlioglu and Ok (2005) and Koeszegi and Rabin (2007). Miguel and I have worked on reference-dependent behaviour, models of sequential choice and stochastic choice models, see, respectively, Apesteguia and Ballester (2009, 2013, 2018) and Apesteguia, Ballester and Lu (2017).

(11) Notable extensions of the work of Afriat are Varian (1990) and Halevy, Persitz and Zrill (2014).

(12) Dean and Martin (2015) refine and extend this index.

(13) Echenique, Lee and Shum (2011) offer a powerful revision of these ideas.

(14) In occasions, the referred papers do not directly report the results described in Table 2. When this is the case, the results of Table 2 are approximations computed by inspecting the datasets used in the different papers.

(15) For other works within this approach see Chambers and Hayashi (2012) and Nishimura (2014).

(16) Other proposals within this second approach are Koeszegi and Rabin (2007), Masatlioglu, Nakajima and Ozbay (2012), Rubinstein and Salant (2012), Caplin, Dean and Martin (2011) and Manzini and Mariotti (2012). See Green and Hojman (2007) for a discussion of the two approaches.

(17) Moreover, the authors suggest a natural refinement of PB that guarantees acyclicity in consumer datasets, and hence is always informative.

References:

Afriat, S. N. (1973), "On a system of inequalities in demand analysis: An extension of the classical method", *International Economic Review*, 14(2), 460-472.

R. Alcantud, J. C., D. L. Matos and C. R. Palmero (2010), "Goodness-of-fit in optimizing a consumer model", *Mathematical and Computer Modelling*, 52(7–8), 1088-1094.

Andreoni, J. and J. Miller (2002), "Giving according to GARP: An experimental test of the consistency of preferences for altruism", *Econometrica*, 70(2), 737-753.

Apesteguia, J. and M. A. Ballester (2009), "A theory of reference-dependent behavior", *Economic Theory*, 40(3), 427-455.

Apesteguia, J. and M. A. Ballester (2010), "The computational complexity of rationalizing behavior", *Journal of Mathematical Economics*, 46(3), 356-363.

Apesteguia, J. and M. A. Ballester (2013), "Choice by sequential procedures", *Games and Economic Behavior*, 77(1), 90-99.

Apesteguia, J. and M. A. Ballester (2015), "A measure of rationality and welfare", *Journal of Political Economy*, 123(6), 1278-1310.

Apesteguia, J. and M. A. Ballester (2018), "Monotone stochastic choice models: The case of risk and time preferences", *Journal of Political Economy,* 126(1), 74-106.

Apesteguia, J., M. A. Ballester. and J. Lu (2017), "Single-crossing random utility models", *Econometrica*, 85(2), 661-674.

Ariely, D. (2008), *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. HarperCollins.

Bernheim, B. D. and A. Rangel (2009), "Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics", *The Quarterly Journal of Economics*, 124(1), 51-104.

Bernheim, B. D., A. Fradkin and I. Popov (2015), "The welfare economics of default options in 401(k) plans", *American Economic Review*, 105(9), 2798-2837.

Blume, L. E. and D. Easley (2008), "Rationality". In S. N. Durlauf and L. E. Blume, (Eds.), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, Basingstoke.

Bouacida, E. and D. Martin (2015), "An application of behavioral welfare economics to consumption data", Working paper.

Caplin, A., M. Dean and D. Martin (2011), "Search and satisficing", *American Economic Review*, 101(7), 2899-2922.

Carvalho, L. S., S. Meier and S. W. Wang (2016), "Poverty and economic decision-making: Evidence from changes in financial resources at payday", *American Economic Review*, 106(2), 260-284.

Chambers, C. P. and F. Echenique (2016), *Revealed Preference Theory*, Cambridge University Press.

Chambers, C. P. and T. Hayashi (2012), "Choice and individual welfare", *Journal of Economic Theory*, 147(5), 1818-1849.

Chetty, R., A. Looney and K. Kroft (2009), "Salience and taxation: Theory and evidence", *American Economic Review*, 99(4), 1145-1177.

Choi, S., S. Kariv, W. Müller. and D. Silverman (2014), "Who is (more) rational? ", *American Economic Review*, 104(6), 1518-1550.

Dean, M. and D. Martin (2015), "Measuring rationality with the minimum cost of revealed preference violations", *Review of Economics and Statistics*, 98(3), 524-534.

Echenique, F., S. Lee and M. Shum (2011), "The money pump as a measure of revealed preference violations", *Journal of Political Economy*, 119(6), 1201-1223.

Einav, L. and L. Yariv (2006), "What's in a surname? The effects of surname initials on academic success", *Journal of Economic Perspectives*, 20(1), 175-187.

Fehr, E. and K. M. Schmidt (1999), "A theory of fairness, competition, and cooperation", *The Quarterly Journal of Economics*, 114(3), 817-868.

Fisman, R., P. Jakiela, S. Kariv and D. Markovits (2015), "The distributional preferences of an elite", *Science*, 349 (6254).

Genesove, D. and C. Mayer (2001), "Loss aversion and seller behavior: Evidence from the housing market", *The Quarterly Journal of Economics*, 116(4), 1233-1260.

Green, J. and D. Hojman (2007), "Choice, rationality and welfare measurement", *Harvard Institute of Economic Research Discussion Paper*, (2144).

Grether, D. M. and C. R. Plott (1979), "Economic theory of choice and the preference reversal phenomenon", *American Economic Review*, 69(4), 623-38.

Halevy, Y., D. Persitz and L. Zrill (2014), "Parametric recoverability of preferences", Technical report, Vancouver School of Economics.

Harbaugh, W. T., K. Krause. and T. R. Berry (2001), "GARP for kids: On the development of rational choice behavior", *American Economic Review*, 91(5), 1539-1545.

Houtman, M. and J. A. H. Maks (1985), "Determining all maximal data subsets consistent with revealed preference", *Kwantitatieve Methoden*, 19, 89-104.

Huber, J., J. W. Payne. and C. Puto (1982), "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis", *Journal of Consumer Research*, 9(1), 90-98.

Itzkowitz, J., J. Itzkowitz and S. Rothbort (2016), "ABCs of trading: Behavioral biases affect stock turnover and value", *Review of Finance*, 20(2), 663-692.

Iyengar, S. S. and M. R. Lepper (2000), "When choice is demotivating: Can one desire too much of a good thing?", *Journal of personality and social psychology*, 79(6), 995-1006.

Kahneman, D. (2011), *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kahneman, D. and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2), 263-291.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1990), "Experimental tests of the endowment effect and the coase theorem", *Journal of Political Economy*, 98(6), 1325-1348.

Köszegi, B. and M. Rabin (2007), "Reference-dependent risk attitudes", *American Economic Review*, 97(4), 1047-1073.

Lacetera, N., D. G. Pope and J. R. Sydnor (2012), "Heuristic thinking and limited attention in the car market", *American Economic Review*, 102(5), 2206-36.

Laibson, D. (1997), "Golden eggs and hyperbolic discounting", *The Quarterly Journal of Economics*, 112(2), 443-478.

Lichtenstein, S. and P. Slovic (1971), "Reversals of preference between bids and choices in gambling decisions", *Journal of Experimental Psychology*, 89(1), 46-55.

Loomes, G., C. Starmer and R. Sugden (1991), "Observing violations of transitivity by experimental methods", *Econometrica*, 59(2), 425-439.

Madrian, B. C. and D. F. Shea (2001), "The power of suggestion: Inertia in 401(k) participation and savings behavior", *The Quarterly Journal of Economics*, 116(4), 1149-1187.

Manzini, P. and M. Mariotti (2012), "Categorize then choose: Boundedly rational choice and welfare", *Journal of the European Economic Association*, 10(5), 1141-1165.

Mas-Colell, A., M. D. Whinston and J. R. Green (1995), *Microeconomic Theory*. Oxford student edition. Oxford University Press.

Mas-Colell, A. (1982), "Revealed Preference after Samuelson". In G. R. Feiwel (Eds.), *Samuelson and Neoclassical Economics*, 72-82. Springer Link, Dordrecht.

Masatlioglu, Y. and E. A. Ok (2005), "Rational choice with status quo bias", *Journal of Economic Theory*, 121(1), 1-29.

Masatlioglu, Y., D. Nakajima and E. Y. Ozbay (2012), "Revealed attention", *American Economic Review*, 102(5), 2183-2205.

May, K. O. (1954), "Intransitivity, utility, and the aggregation of preference patterns", *Econometrica*, 22(1), 1-13.

Mullainathan, S. and R. Thaler (2001), "Behavioral Economics". In N. Smelser and P. Bates (Eds.), *International Encyclopedia of Social Sciences*, 1094-1100. Pergamon Press.

Nasar, S. (2011), *Grand Pursuit: The Story of Economic Genius*. Simon & Schuster.

Nishimura, H. (2014), "The transitive core: Inference of welfare from nontransitive preference relations", working paper 201419, University of California at Riverside.

Rabin, M. (1993), "Incorporating fairness into game theory and economics", *The American Economic Review*, 83(5), 1281–1302.

Roelofsma, P. H. M. P. and D. Read (2000), "Intransitive intertemporal choice", *Journal of Behavioral Decision Making*, 13(2), 161-177.

Rubinstein, A. (2012), *Lecture Notes in Microeconomic Theory: The Economic Agent*, Princeton University Press.

Rubinstein, A. and Y. Salant (2012), "Eliciting welfare preferences from behavioural data sets", *The Review of Economic Studies*, 79(1), 375-387.

Samuelson, P. A. (1938), "A note on the pure theory of consumer's behaviour", *Economica*, 5 (17), 61-71.

Simonson, I. (1989), "Choice based on reasons: The case of attraction and compromise effects", *Journal of Consumer Research*, 16(2), 158-174.

Smeulders, B., F. C. R. Spieksma, L. Cherchye and B. De Rock (2014), "Goodness-of-fit measures for revealed preference tests: Complexity results and algorithms", *ACM Transactions on Economics and Computation*, 2(1), 3-16.

Swofford, J. L. and G. A. Whitney (1986), "Flexible functional forms and the utility approach to the demand for money: A nonparametric analysis: Note", *Journal of Money, Credit and Banking*, 18(3), 383-389.

Thaler, R. H. (2015), *Misbehaving: The Making of Behavioral Economics*. W.W. Norton.

Thaler, R. H. (1999), "Mental accounting matters", *Journal of Behavioral Decision Making*, 12(3), 183-206.

Tversky, A. and D. Kahneman (1981), "The framing of decisions and the psychology of choice", *Science*, 211(4481), 453-458.

Tversky, A. (1969), "Intransitivity of preferences", *Psychological Review*, 76(1), 31-48.

Tversky, A. and R. H. Thaler (1990), "Anomalies: Preference reversals", *Journal of Economic Perspectives*, 4(2), 201-211.

Varian, H. R. (1990), "Goodness-of-fit in optimizing models", *Journal of Econometrics*, 46(1-2), 125-140.

Opuscles already published

1. **Reconsidering Spanish unemployment**
   Ramon Marimon (June 97)
2. **Reducing unemployment. At any cost?**
   Fabrizio Zilibotti (December 97)
3. **Capital and labor taxes, macroeconomic activity, and redistribution**
   Albert Marcet (November 98)
4. **The lender of last resort in today's financial environment**
   Xavier Freixas (November 99)
5. **Why does the public sector grow? The role of economic development, trade and democracy**
   Carles Boix (November 99)
6. **Gerontocracy and social security**
   Xavier Sala-i-Martin (July 2000)
7. **The political viability of labour market reform**
   Gilles Saint-Paul (December 2000)
8. **Are EU policies fostering growth and reducing regional inequalities?**
   Fabio Canova (May 2001)
9. **Agglomeration effects in Europe and the USA**
   Antonio Ciccone (September 2001)
10. **Economic polarization in the Mediterranean basin**
    Joan Esteban (May 2002)
11. **How do households invest their wealth?**
    Miquel Faig (October 2002)
12. **Macroeconomic and distributional effects of social security**
    Luisa Fuster (April 2003)
13. **Educating intuition: A challenge for the 21st century**
    Robin M. Hogarth (September 2003)
14. **Capital controls in post-war Europe**
    Hans-Joachim Voth (April 2004)
15. **Taxation of financial intermediaries**
    Ramon Caminal (September 2004)
16. **Ready to take risks? Experimental evidence on risk aversion and attraction**
    Antoni Bosch-Domènech / Joaquim Silvestre i Benach (November 2005)
17. **Social networks and labour market outcomes**
    Antoni Calvó-Armengol (January 2006)
18. **The effects of employment protection in Europe and the USA**
    Adriana D. Kugler (February 2007)

19. **Urban sprawl: Causes and consequences**
Diego Puga (January 2008)

20. **Western European long term growth, 1830-2000: Facts and issues**
Albert Carreras and Xavier Tafunell (June 2008)

21. **Overcoming coordination failure in firms and organizations: Experimental evidence**
Jordi Brandts (March 2009)

22. **The misallocation of talent**
José V. Rodríguez Mora (May 2009)

23. **Complementarities in innovation strategy and the link to science**
Bruno Cassiman (September 2009)

24. **Simple mechanisms to resolve conflicting interests and to share the gains**
David Pérez-Castrillo (November 2009)

25. **Transfer of university innovations**
Inés Macho-Stadler (January 2010)

26. **Firing costs, dismissal conflicts and labour market outcomes**
Maia Güell (June 2010)

27. **Inequality and tax progressivity**
Juan Carlos Conesa (October 2010)

28. **Happiness economics**
Ada Ferrer-i-Carbonell (May 2011)

29. **School choice in Spain: Theory and evidence**
Caterina Calsamiglia (September 2011)

30. **Corporate social responsibility and social welfare. How to promote socially responsible business strategies**
Juan-José Ganuza (March 2012)

31. **The effectiveness of foreign aid: From the macroeconomic perspective to the experimental evaluation**
José G. Montalvo / Marta Reynal-Querol (June2012)

32. **Fiscal policy in the European Monetary Union**
Evi Pappa (September 2012)

33. **How important are capital markets Imperfections in determining firm decisions and aggregate fluctuations?**
Andrea Caggese (November 2012)

34. **Globalization, technology and inequality**
Gino Gancia (April 2013)

35. **Credit cycles and systemic risk**
José-Luis Peydró (December 2013)

36. **The labour market impact of immigration**
Albrecht Glitz (June 2014)

37. **Density forecasts in economics and policymaking**
Barbara Rossi (September 2014)

38. **Local labor markets**
Jan Eeckhout (December 2014)

39. **Globalization and aggregate fluctuations: The role of international trade and large firms**
Julian di Giovanni (February 2015)

40. **The social effects of divorce laws**
Libertad González (June 2015)

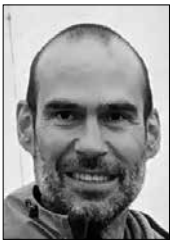
41. **Why do people buy private health insurance in the presence of a public option?**
Pau Olivella (July 2016)

42. **Health interventions in low income countries: A (not so) low hanging fruit?**
Alessandro Tarozzi (October 2016)

43. **Media failures and media markets**
Fabrizio Germano (January 2017)

44. **Mass media and its influence on behaviour**
Ruben Enikolopov and Maria Petrova (December 2017)

45. **Measures of rationality and welfare for behavioural decision-makers**
Jose Apesteguia (July 2018)



Jose Apesteguia

Jose Apesteguia is Professor of Economics at the Universitat Pompeu Fabra, Research Professor at ICREA, and Barcelona GSE Research Professor. Jose received his PhD in Economics from the Public University of Navarra in 2001. He studies individual decision-making with an emphasis on its psychological foundations. He uses theoretical tools as well as adopts an empirical approach to try to better understand individual decision-making, and to improve the predictive power of decision-making models in economic environments. Jose Apesteguia has published his research in leading international journals such as *American Economic Review, Econometrica, Journal of Political Economy and Review of Economic Studies.*

CREi
Centre de Recerca
en Economia Internacional

Academic Unit of the Barcelona GSE Graduate School of Economics

Ramon Trias Fargas, 25-27 - 08005 Barcelona
Tel: 93 542 13 88 - Fax: 93 542 28 26
E-mail: crei@crei.cat
http://www.crei.cat

upf. Universitat Pompeu Fabra Barcelona

Generalitat de Catalunya