**Assessing Axioms of Theories of Limited Attention**

Nuttaporn Rochanahastin

Department of Economics and Related Studies, University of York, York, YO10 4GA, United Kingdom

E-mail: nr708@york.ac.uk

**Abstract**

**Assessing Axioms of Theories of Limited Attention**

It is very likely that many decisions in the individual choice setting are made without a complete or exhaustive deliberation process. One of the plausible explanations for this behaviour, that has recently received much recognition in economics, is that people have limited attention. This assumption has motivated many new theories, more and more of which are founded on axioms. This research experimentally test two of these new theories, those of Masatlioglu *et al.* (2012), and Lleras *et al.* (2017), which are based on the revealed preference framework. This paper uses standard choice data to determine the (relative) validity of their underlying axioms, compared to a benchmark of the violation rate given by random behaviour. The results show that Masatlioglu *et al.* (2012) appears to be the empirically more plausible weakening of WARP.

Keywords: Experiments, Limited attention, Revealed preferences, Axioms

JEL classifications: C91, D01, D12, D83, D91

**1. Introduction**

Most economic theories are built upon axioms. This is particularly true for decision theories and social choice theories. The validity of the predictions coming out of these theories depends upon the validity of the underlying axioms. In a strict sense an axiom can only be right or wrong: one observation violating an axiom can be considered proof that it is wrong. A cynic would argue that all axioms are wrong, and, while that is almost certainly true of axioms in economics, it is not particularly helpful. One way of rationalising violations is to posit that decision-makers make ‘occasional’ mistakes – that is, there is some noise in their behaviour. We then need to find a way of measuring the amount of noise in behaviour (relative to the theory being tested). We need to measure ‘how right’ an axiom is.

Research on a new batch of theories addressed to *satisficing*, or sub-optimality,behaviour stimulated the direct test of axioms as most of them are axiomatic based, but the testing methodology can also be applied in other contexts. The word *satisficing* was coined by Simon (1955) as describing behaviour which is not optimising ‒ behaviour in which the decision-maker aims for a satisfactory outcome rather than an optimal one. There are many theories which try and describe satisficing behaviour, including models of incomplete preferences, models of behaviour under ambiguity, theories of rational inattention, and search theories.

An area which is particularly active is that of theories of rational inattention. This current paper focuses on this area. This area has recently caught the attention of researchers following the pioneering works of Sims (1998, 2003, 2010). The applications of this body of research extend to wide areas such as Macroeconomics, Games, etc.[[1]](#footnote-1) Here we focus on the works that use the standard choice data and Revealed Preference methods.

Manzini and Mariotti (2007), Masatlioglu *et al.* (2012), and Lleras *et al.* (2017) all have the same structure: a decision-maker (DM) is being asked to choose one element out of some large choice set, but the set is so large that the DM, in order to simplify a complex problem, pays attention to, and hence chooses from, a subset of this set – a subset called the *Consideration Set*. Axioms characterise how the DM does this. The similarity of these three papers is that all (have to) weaken a standard axiom of decision theory, namely the Weak Axiom of Revealed Preference (WARP). They do it in different ways with different weakenings. This sub-branch of rational inattention theories is appropriate with standard choice data and in line with the experimental design use in this research.

Manzini and Mariotti (2007) provides a ‘shortlisting method’. That paper suggest a two-stage procedure, in which the DM in the first stage weeds out unacceptable choices using one criterion and then proceeds in the second stage to a choice using another criterion. Such procedure is called a ‘Rational Shortlist Method’ (RSM). They later, in 2012, suggested a relaxation to RSM called ‘Categorise Then Choose’ (CTC) which allows the DM to compare sets of alternatives in the first stage. They suggest that the theories are testable by testing axioms such as WARP, weakening of WARP, and the Expansion axiom (An alternative chosen from each of two sets is also chosen from their union). Cherepanov *et al.* (2013) suggest a similar procedure to CTC in that the DM compares several rationales or ‘motivations’ in the first stage and maximises preference among shortlisted alternatives in the second stage. Manzini and Mariotti (2010) experimentally tested some of these theories; their results are reported in the next section.

**2. Theories being tested and relevant literatures**

Masatlioglu *et al.* (2012) suggests the use of an ‘attention filter’[[2]](#footnote-2) to be the main property of, and which determines, a member of a consideration set. In essence, the attention filter requires that a consideration set is unaffected when an alternative that a DM does not pay attention to becomes unavailable. As a result of this property, the DM is revealed to pay attention to some alternatives. A choice reversal is needed in order to elicit the DM’s preference. This is a direct contradiction of WARP. The model is empirically testable by testing the axiom of WARP with Limited Attention (WARP(LA)) which is:

For any nonempty S, there exists x\* ∈ S such that, for any T including x\*

if c(T ) ∈ S and c(T ) ≠ c(T \x\* ), then c(T ) = x\*.

According to this model, *x* is revealed preferred to *y* if and only if when *y* is taken out of the choice set, *x* is no longer chosen. *y* is revealed to attract attention from the DM while *x* is chosen in the original set. This axiom provides an interesting and crucial implication which is the acyclicity property. Preferences implied using the theory is acyclic. An example of a cycle preference relation would be: $a\_{1}≻a\_{2}≻…≻a\_{k}≻a\_{1}$.

Lleras *et al.* (2017) also points out that a consideration set and its primitives might not be directly observable. However, their property is empirically testable. The paper based its theory on the assumption of contraction consistency[[3]](#footnote-3) and coined the term ‘competition filter’ to be the main property of a consideration set, and subsequently, the revealed preference. The paper’s main axiom is Limited Consideration WARP (LC-WARP):

 For any nonempty S, there exists b\*∈ S such that for any T including b\*,

 if (i) c(T) ∈ S, and

(ii) b\* = c(T’) for some T’ $⊃$ T

then c(T) = b\*

Again, this axiom’s main implication is that it does not allow any cycle in a choice function. Masatlioglu *et al.* (2012) and this paper are the main focus of this research. These two papers are one of the first papers in the rational inattention field, to the author knowledge, that do not impose unobservable restrictions on the consideration set[[4]](#footnote-4) and have not been experimentally tested. Their characterizations are also based on WARP which is empirically testable from directly observed choice functions.

There are also other related models that involve two-stage shortlisted procedures. However, they are not focused on experimentally in this paper because the models impose some requirements or assumptions on consideration set formation or the shortlisting procedures which make them not compatible with the experimental design in this paper[[5]](#footnote-5). Manzini *et al.* (2013) provides characterisation for Two-Stage Threshold representation (TST). Alternatives survive the first stage screening if a threshold value is reached. In a series of papers Tyson (2008, 2013, 2015) developed extensively two-stage incomplete preference models which provide some connections between satisficing to salience and attention. Finally, search and costly information acquisition was developed in Caplin *et al.* (2011), Caplin and Dean (2011, 2015), and Matějka and McKay (2015). Manzini and Mariotti (2014) provides a probabilistic version of a consideration set. A random consideration set is a randomly drawn subset of the choice set. The actual choice is the most preferred item in the consideration set.

In terms of empirical literatures, Manzini and Mariotti (2010) is the closest in spirit to this research. They report on a choice experiment using remuneration instalments as alternatives. There are 4 instalment plans and subjects were presented with all combinations of them. The paper investigates axioms from standard decision theory (WARP) as well as other three theories, namely RSM, CTC and Cherepanov *et al.* (2013)’s version of Rationalisation. They find that one aspect of WARP (Condorcet) is violated substantially more than the other (pairwise choice). Therefore, models that are more compatible with the Condorcet property, for example, CTC, are more likely to be successful in the experiment. As expected, WARP is violated the most. On the other hand, CTC and Order Rationalisation perform well. They also use Selten’s Measure of Predictive Success, which introduces a parsimonious factor, to take into consideration some of the nested-ness of these models.

Chetty *et al.* (2009) observed inattention in the case of taxation. They conducted a field experiment observing the difference between tax-inclusive and tax-exclusive price tags in a grocery store and find that changes in tax policy affect demands more in tax-inclusive price tags. De los Santos *et al.* (2012) uses data on web browsing and online book purchasing to test search models. The paper rejects a sequential search model in which a consumer always buys from the last store she visited, when he/she crosses the reservation benchmark and favours the fixed sample size search strategy.

Caplin *et al.* (2011) report on a search experiment within which there were four ‘Experiments’. In Experiment 1 search was over a set of payoffs expressed as simple sums (“two plus eight minus six”) differing in their number and complexity, with no time constraint; the sums were generated from an exponential distribution (shown to the subjects). In Experiment 2, subjects were told that their payment will be at a random time in a decision problem. This is to incentivise subjects to always choose the best alternative at that moment in time. Experiment 3 was designed to explore how screen position and object complexity impacts search order. Experiment 4 was the same as Experiment 1 with a two-minute time constraint. The novelty of this paper is that it recorded provisional choice data and contemplation times. They find evidence supporting the sequential search and satisficing model.

**3. Experimental Design**

The purpose of this research is to experimentally determine which of Masatlioglu *et al.* (2012) or Lleras *et al.* (2017)) appear to be the empirically more plausible WARP weakenings, if at all empirically better than WARP itself. Therefore, the experimental procedure is relatively close to Manzini and Mariotti (2010). It is a choice-function-eliciting experiment where the alternatives are risky lotteries. However, this is an attention-related experiment, thus the more alternatives presented to the subject the better. This gives higher chance of preventing subjects from recognizing the pattern of the alternatives or carefully deliberating through each of the problem. The drawback is that there is a limitation on the number of problems that the experimenter can present to the subjects. Therefore, we presented only subsets of problems.

This experiment had 10 baseline lotteries[[6]](#footnote-6) which imply a total of 1,023 possible subsets. The subjects were presented with 118 of them[[7]](#footnote-7). The lotteries are designed to be similar but contain some difference in features. The expected value of lotteries vary from a minimum of £8.00 to a maximum of £9.80. The randomised process in selecting the subsets starts from randomly selecting a subset of two alternatives and based on that randomly selecting a higher number of alternatives’ subsets. We, first, randomly selected 5 2-alternative subsets and based on that, we randomly selected 3-alternative subsets that are supersets of one of those 5 2-alternatives subsets. After that, we randomly selected 4-alternative subsets that are supersets of one of those 3-alternatives subsets, and so on. The lottery visualisation is in two-dimensional figure where the x-axis represent probabilities and y-axis represent money outcomes. Because there are two important attributes that comprised a lottery, money outcome and probability. We feel that two-dimensional figure best capture this concept as well as give subject some ideas of the expected value of a lottery in the form of shaded areas. Figure 1 shows an example of a lottery and how it is presented to the subjects. This example lottery has 8 in 10 chance of gaining £7 and 2 in 10 chance of gaining £11.



Figure 1. A visualisation of a lottery.

Figure 2 shows an example screenshot of a problem faced by the subjects. The subject’s task is relatively straightforward and simple which is to choose the most preferred lottery in each problem. Again, taking into account that this is an attention experiment, an upper bound of 45 seconds per problem was imposed. Subjects had to wait a minimum of 10 seconds before confirming their choice, to minimise them clicking at random.



Figure 2. An experimental screenshot.

At the end of the experiment, for each subject, the chosen lottery in a randomly selected problem was played out for real. Each subject randomly selected a problem for their payment in private by drawing a disk from a bag containing numbered disks from 1 to 118. Their lottery choice in that problem was then played out for real by drawing from another bag containing 10 disks, a multiple of 10 from 10 to 100. The total payment for the experiment is the lottery payoff plus £3 show-up fee. Subjects were informed that some lotteries involve losses. The maximum loss outcome of any lottery is £3. If subject’s lottery payoff is a loss then this is taken out from the show-up fee. After the payment, subjects were free to go.

We recruited a total of 65 subjects for the experiment which was conducted in the EXEC Lab at the University of York. Subject’s ages ranged from 18 to 44 years. 64 of whom were students and one was a member of staff at the University of York. There were 34 females (52.31%) and 31 males (47.69%). The average total payment per subject was £11.25. Subjects spent an average of less than one hour in the laboratory. This experiment was run using purpose-written software written in Visual Studio.

**4. Results**

The Weak Axiom of Revealed Preference (WARP) is the textbook normative axiom and a baseline description of utility maximisation behaviour. Choice inconsistencies violate the axiom which is essential to the utility maximisation model. Preference reversals and choice inconsistencies have long been confirmed by a large number of literatures[[8]](#footnote-8). Most of these literatures use observed individual choice because it is the most obvious and normatively appealing as a measure of preference. Therefore, preference from observed choices plays an important role in this analysis and preference reversals serve as a key measurement of axiom violations in the various models.

The preferences will be extracted from choice(s) given each model’s requirement. First, the preference inference for WARP is direct and straight forward. The axiom states that, in every choice set, there is the best alternative that must be chosen. It means that the chosen alternative from a choice set is revealed preferred to the other alternatives in the set. Therefore, for every problem, pairwise preference(s) can be inferred. Clearly, the axiom has an implicit assumption of full attention, namely that a DM considers every alternative in the choice set. For Masatlioglu *et al.* (2012), limited attention consideration is taken into account for the inferred preferences. We need to make sure that the alternative attracts attention in order to be able to extract a preference. An alternative *x* is revealed preferred to *y* if and only if when *y* is taken out of the choice set, *x* is no longer chosen. For example, if a DM choses *a1* from a choice set {*a1, a2, a3*} and *a3* from a choice set {*a1, a3*}, we can conclude that *a1* is revealed preferred to *a2* because dropping *a2* changes the choice which means that DM must have paid attention to *a2* when he/she chose from {*a1, a2, a3*}. Notice that this is a direct contradiction to WARP. WARP needs to be violated in order for Masatlioglu *et al*. model to infer something. Also, we need at least two problems and a choice reversal to be possible to infer any preference.

Lleras *et al.* (2017) based their consideration set formation under the assumption that if an alternative attracts the DM in the menu with more alternatives, it will also grab his/her attention in subsets of the menu. A choice change in a smaller menu suggests that the choice is preferred to that from the bigger menu that is its superset. For example, if a DM chose *a1* from a choice set {*a1, a2, a3*} and *a3* from a choice set {*a1, a3*}, we can conclude that *a3* is revealed preferred to *a1*because the DM must have seen *a1* from the choice set {*a1, a3*}. This theory is also a direct contradiction to WARP and there is a possibility that preferences inferred are incomplete. These three models uncover preferences from observed choices under different (and contradictory) assumptions. This section analyses the relative strength in term of explanatory power of each axiom. The analysis begins by examining how complete. Then, the inconsistencies in various aspects are analysed.

To make it clearer, let us provide an example of how preferences can be inferred from each theory given the experimental design. For simplicity and without the loss of generality, suppose there are 4 alternatives *a1, a2, a3,* and *a4* and the details of problems given to subjects are given in columns 1 and 2 of table 1. Suppose that the DM’s decisions are as column 3.

|  |  |  |
| --- | --- | --- |
| Problem No. | Choice set | Decision |
| 1 | {*a1, a2, a3, a4,*} | *a3* |
| 2 | {*a1, a2, a3*} | *a1* |
| 3 | {*a1, a2, a4*} | *a2* |
| 4 | {*a1, a3, a4*} | *a1* |
| 5 | {*a2, a3, a4*} | *a4* |
| 6 | {*a1, a2*} | *a1* |
| 7 | {*a1, a3*} | *a3* |
| 8 | {*a1, a4*} | *a4* |
| 9 | {*a2, a3*} | *a2* |
| 10 | {*a2, a4*} | *a2* |
| 11 | {*a3, a4*} | *a4* |

 Table 1: Examples of problems and decision.

Table 2 reports all the *pairwise* preferences that can be inferred from table 1[[9]](#footnote-9). The first column of table 2 is all the permutation pairs from a set of 4 alternatives. By [*a,b*]*[[10]](#footnote-10)* we mean that *a* is preferred *b*. For an example of how to infer preferences according to WARP, let us take a look at problem number 1. *a3* is chosen, therefore, we can infer [*a3, a1*], [*a3, a2*], and [*a3, a4*]. Applying the same process to other problems give us the second column of the table. Masatlioglu *et al*. requires at least two problems to infer any preference. Let us take a look at the first and the second problems: we note that dropping *a4* changes the choice; therefore, we can conclude that *a4* must attract attention of the DM in problem 1 but *a3* is chosen. Hence, *a3* is revealed preferred to *a2*. Again, apply the same process to every pairs of problem give us the third column of table 2. The inference for Lleras *et al* also requires at least two problems. Again, let us take a look at the first two problems: *a3* is chosen in the first problem suggesting that *a3* must attract DM attention in every smaller subset that contains *a3*. Hence, we can infer [*a1, a3*] because *a1* is chosen in the second problem.

|  |  |  |  |
| --- | --- | --- | --- |
| Pairwise Permutations | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| [*a1, a2*] | ✗ | ✗ | ✗ |
| [*a1, a3*] | ✗ |  | ✗ |
| [*a1, a4*] | ✗ | ✗ |  |
| [*a2, a3*] | ✗ |  | ✗ |
| [*a2, a4*] | ✗ | ✗ | ✗ |
| [*a3, a4*] | ✗ | ✗ |  |
| [*a2, a1*] |  |  |  |
| [*a3, a1*] | ✗ | ✗ | ✗ |
| [*a4, a1*] | ✗ |  | ✗ |
| [*a3, a2*] | ✗ | ✗ |  |
| [*a4, a2*] | ✗ |  |  |
| [*a4, a3*] |  | ✗ | ✗ |

Table 2: All inferred pairwise preference.

It is noticeable that there are conflicting pairwise preference relationships in table 2. For example, both [*a1, a3*] and [*a3, a1*] are checked in WARP column. These represent cycles which are violations to the axioms. If we extract only non-conflicting or acyclic pairwise preference, we can fill table 3 which reports only the pairwise combinations of the alternatives. The marks in the third to the fifth column show that we can infer the valid cycle-free[[11]](#footnote-11) relationships within the pair. For example, (*a1, a2*)[[12]](#footnote-12) can be either *a1* $≻$ *a2 or a1* $≺$*a2*.

|  |  |  |  |
| --- | --- | --- | --- |
| Pairwise Combinations | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| (*a1, a2*) | ✗ | ✗ | ✗ |
| (*a1, a3*) |  | ✗ |  |
| (*a1, a4*) |  | ✗ | ✗ |
| (*a2, a3*) |  | ✗ | ✗ |
| (*a2, a4*) |  | ✗ | ✗ |
| (*a3, a4*) |  |  | ✗ |

Table 3: Valid inferred pairwise preference.

**4.1 Inferred preferences**

First, we take a look at how much (or how complete) preferences each theory can infer, given the same set of problems (choice sets). All these theories can infer direct pairwise preferences according to the methods mentioned above. In term of permutations, there are 90 possible pairwise preference relationships which can in principle be inferred from the experimental data[[13]](#footnote-13). The inference percentage per subject is calculated out of these 90 relationships. Table below reports the average percentages over all subjects. The simulation’s method and procedure will be explained after the table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| All inferred pairwise preferences | Actual | Simulation | Absolute Difference | Relative Difference |
| WARP | 69.62% | 96.32% | -26.70 p.p.[[14]](#footnote-14) | -38.35% |
| Masatlioglu et al. | 48.02% | 63.15% | -15.13 p.p. | -31.51% |
| Lleras et al. | 62.32% | 41.33% | 20.99 p.p. | 33.68% |

Table 4: All inferred pairwise preferences.

Since crudely comparing these numbers to determine the relative validity of axioms will not work, we have developed a method of providing a ‘benchmark’. This is derived from simulations of random behaviour and counting the number of inferred preferences. At the end of the day we can compare the observed number of violations of each axiom with the benchmark figures and hence provide a relative measurement of ‘how good’ is each axiom. This also penalises for different degrees of restriction of each theory to give a fairer competing ground in the comparisons. In this case, the simulation is done by creating 100,000 repetition of random decisions using the same 118 problems that subjects faced. These decisions are used to extract preferences in the same manner as the actual data (the same procedures explained in the example given in table 2 and 3 above). This method can serve as one of the suggestions for a direct and non-parametric test of these theories.

The *t*-test for difference between two population means is employed to verify that the average from the actual experimental data is significantly different from the average from the simulations for each theory. We found that the *p*-value for all three theories are equal to 0.0000, suggesting that the two means are significantly different. WARP is more restricted than the other two models so the random behaviour provide the highest inferred preference percentage at 96.32%. This is followed by Masatlioglu *et al*. at 63.15% and the then by Lleras *et al* at 41.33%. The actual data shows that WARP declines the most in term of absolute percentage points, from 96.32% to 69.62%, but remains the highest in term of the ability to extract preferences which suggests that it is the most restricted model. In terms of relative differences, WARP and Masatlioglu *et al*. are both decline over 30% relative to simulation while Lleras *et al*. improves over 30%. Lleras *et al*. is shown to be more restricted than anticipated from the random behaviour while Masatlioglu *et al*. is the least restricted model.

**4.2 The inconsistencies**

Next, the cyclicity of each theory is analysed in different aspects. Choice inconsistencies or revealed preference cycles are the main criteria that can be used to measure the relative degree of validity in the three theories as all three characterisations involve a common acyclicity property. We are going to look at the breadth, depth, and length of the cycles based on the categorisation by Bouacida and Martin (2017). The breadth of the cycles are how spread cycles are observed among experimental subjects[[15]](#footnote-15). The depth and length of the cycles delve deeper into individual behaviour. The depth investigates direct inferred pairwise preferences while the length applies transitivity assumption.

**4.2.1 The breadth of the cycles**

First, we want to look at how widely spread the inconsistencies are shown among experimental subjects. The data revealed that the inconsistencies are much more extensive in this experiment (10 alternatives) as compared to Manzini and Mariotti (2010) (4 alternatives). We found that 100% of the subjects shown some degree of choice inconsistencies according to WARP and Lleras *et al*. while only 6.15% (4 out of 65 subjects) display consistent preferences according to Masatlioglu *et al*. This finding is as expected and consistent with empirical literatures observing pervasive preference cycles in choice behaviour.

**4.2.2 The depth of the cycles**

Next, we take a look at how much cycles invaded into the inferred preference. We begin by analysing the all inferred pairwise preference (as presented in table 5). The depth of the cycles are represented by the proportion of those inferred preference that exhibit inconsistencies. This can be calculated by dividing the number of inferred pairs that exhibit cycles by the total number of pairs inferred. Table 5 reports these percentages. The hypothesis is that the higher the percentage of the cycles, the more violation of the axioms are shown in the data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Depth of the cycles | Actual | Simulation | Absolute Difference | Relative Difference |
| WARP | 58.73% | 96.70% | -37.97 p.p. | -64.65% |
| Masatlioglu *et al*. | 34.11% | 58.47% | -24.36 p.p. | -71.42% |
| Lleras *et al*. | 65.94% | 55.76% | 10.18 p.p. | 15.44% |

Table 5: Depth of the cycle.

Because the denominator in the calculations is the inferred preferences, these percentages are already take into consideration the degree of restriction. Again, we use the *t*-test for difference between two population means and found that the *p*-value for all three theories are equal to 0.0000, rejecting that the null hypotheses of equal means. Masatlioglu *et al*. shows the greatest improvement, in term of the relative difference, compared to the simulations while the Lleras *et al*. violation percentage increases relative to the simulation. Masatlioglu *et al*. also has the lowest actual violation percentage while Lleras *et al*. is the highest. WARP shows a significant improvement from the simulation that displays almost 100% violation rate and the actual violation percentage is still relatively higher than Masatlioglu *et al*.

Next, we delve deeper into the validity of each axiom by focusing on the valid (consistent) inferred preference. This can be done by observing the inferred preferences calculated in section 4.1 and extracting only those pairwise preference combinations that are not exhibit any inconsistency[[16]](#footnote-16). The proportion of these valid relations over the total number of pairwise choice combinations (45 pairs) are calculated, and reported in the second column of table 6 – Valid inferred pairwise preference. This also shows how complete of the inferred preference, taken into the account only consistent preference. The hypothesis here is the higher the valid inferred preference, the more consistent and complete is the axiom.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Valid inferred preference | Actual | Simulation | Absolute Difference | Relative Difference |
| WARP | 50.97% | 0.63% | 44.71 p.p. | 87.72% |
| Masatlioglu *et al*. | 60.41% | 52.07% | 8.34 p.p. | 13.81% |
| Lleras *et al*. | 36.38% | 36.27% | 0.11 p.p. | 0.30% |

Table 6: Valid inferred pairwise preferences.

Intuitively, the valid inferred pairwise preferences are inversely related with the all inferred preferences. More restricted models result in higher preference inference, which in turn, translate into higher chance of cycles and less valid inferred preferences. The *t*-test of difference in means rejects the null hypothesis for WARP and Masatlioglu *et al*. while it fails to reject the null hypothesis for Lleras *et al*. with a *p*-value of 0.4835. The results show that Masatlioglu *et al*. has the highest percentage of valid inferred preferences which is as expected because it is the least restricted model. It shows an improvement of 13.81% relative to the simulation. However, WARP has the highest increment on the relative difference. It is an improvement relative to only 0.63% of the inferred preference in simulation because it is the most restricted model. This shows modest support to both Masatlioglu *et al*. and WARP. Lleras *et al*. does not do well in this category. The valid inferred preference is only 36.38%. Its validity and completeness are only marginally improved over the simulation of random behaviour and in fact, is significantly indifferent. A plausible explanation is that Lleras *et al*. is more restrictive compared to Masatlioglu *et al*. as observed from higher percentage of all inferred pairwise preference (table 4). However, using the same criterion, WARP is the most restricted model but still perform better than Lleras *et al*. in both analyses and in the depth-of-cycles category.

**4.2.3 The length of the cycles**

The violations in each cycle length can be obtained by observing the violations assuming transitive preference. The shortest possible length here is a cycle of length 2. This is a direct inconsistency or a reverse in preference. This type of cycles is already analysed in the previous section. Longer lengths are obtained from applying the transitive assumption and the longest length is 10. We calculate the violation percentage[[17]](#footnote-17) in each length. It is unclear whether the longer or the shorter the length is more problematic for the axioms. One can argue that cycle of length 2 is the direct contradiction but also on the other hand, it is not sensible for the longer transitive preference to have a contradiction as the preference ranking should be clearer.

|  |  |  |  |
| --- | --- | --- | --- |
| Cycle Length | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| 2 | 58.73% | 34.11% | 65.94% |
| 3 | 75.48% | 56.47% | 83.34% |
| 4 | 79.81% | 67.75% | 87.99% |
| 5 | 80.18% | 72.14% | 88.37% |
| 6 | 80.18% | 72.14% | 88.37% |
| 7 | 80.18% | 72.14% | 88.37% |
| 8 | 80.18% | 72.14% | 88.37% |
| 9 | 80.18% | 72.14% | 88.37% |
| 10 | 80.18% | 72.14% | 88.37% |

Table 7: Violations in different cycle length from actual data.

|  |  |  |  |
| --- | --- | --- | --- |
| Cycle Length | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| 2 | -37.97 p.p. | -24.36 p.p. | 10.18 p.p. |
| 3 | -24.40 p.p. | -38.27 p.p. | -0.59 p.p. |
| 4 | -20.07 p.p. | -30.79 p.p. | -7.19 p.p. |
| 5 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |
| 6 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |
| 7 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |
| 8 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |
| 9 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |
| 10 | -19.70 p.p. | -26.40 p.p. | -6.92 p.p. |

Table 9: Absolute difference of violations in cycle lengths.

|  |  |  |  |
| --- | --- | --- | --- |
| Cycle Length | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| 2 | 96.70% | 58.47% | 55.76% |
| 3 | 99.88% | 94.74% | 83.93% |
| 4 | 99.88% | 98.54% | 95.19% |
| 5 | 99.88% | 98.54% | 95.29% |
| 6 | 99.88% | 98.54% | 95.29% |
| 7 | 99.88% | 98.54% | 95.29% |
| 8 | 99.88% | 98.54% | 95.29% |
| 9 | 99.88% | 98.54% | 95.29% |
| 10 | 99.88% | 98.54% | 95.29% |

Table 8: Violations in different cycle length from simulations.

|  |  |  |  |
| --- | --- | --- | --- |
| Cycle Length | WARP | Masatlioglu *et al*. | Lleras *et al*. |
| 2 | -64.65% | -71.42% | 15.44% |
| 3 | -32.32% | -67.76% | -0.07% |
| 4 | -25.15% | -45.44% | -8.18% |
| 5 | -24.57% | -36.60% | -7.84% |
| 6 | -24.57% | -36.60% | -7.84% |
| 7 | -24.57% | -36.60% | -7.84% |
| 8 | -24.57% | -36.60% | -7.84% |
| 9 | -24.57% | -36.60% | -7.84% |
| 10 | -24.57% | -36.60% | -7.84% |

Table 10: Relative difference of violations in cycle lengths.

The percentage increases upon each length because of the transitivity assumption. Table 7 reports the results of cycle lengths from the experimental data while table 8 reports results from the simulations. Tables 9 and 10 report their differences. The differences in means are significant at every length for WARP and Masatlioglu *et al*. while the differences for Lleras *et al*. are significant at every length but length 3. We can see that the pattern remains through every length in term of actual violation. Masatlioglu *et al*. shows lowest violation rates in every length while Lleras *et al*. shows the highest. The maximum violations for Masatlioglu *et al*. is 72.14% compare to WARP at 80.54% and Lleras *et al*. at 88.37%. Both Masatlioglu *et al*. and WARP improve at every cycle length when compare to the simulations. However, Lleras *et al*. performs worse than simulations for cycles of length 2. Lleras *et al*. did improve over simulation when transitivity is assumed in the longer lengths. In term of relative difference, Masatlioglu *et al*. still shows greatest improvement over simulations when transitivity is fully explored, follows by WARP and Lleras *et al*.

Since these three theories provide different predictions and contain overlapping areas. One might argue that there is a need to penalise in order to compare their explanatory power. We tried to address this issue by using the ‘benchmark’ procedure. One possible alternative method is Selten’s measure of predictive success (Selten 1991). The measure is given by:

$$m=r-a$$

where $r$ is the relative frequency of correct prediction (the number of observed outcomes divided by the number of possible outcomes). While $a$ is the penalised parameter which is given by the size of the predicted subset compared with the set of all possible outcomes.

There is a practical difficulty of this measure namely, the number of outcomes increases drastically with the number of the number of problems or alternatives. For this study, predictive parsimony variable ($a$) in the measure for WARP, given there are 10 alternatives and 118 problems is (3628800/3\*10^74) which is approximately zero. The $a$ variable is also the same (zero) for Masatlioglu *et al*. and Lleras *et al*. since the denominator is also very large. Therefore, the measure is left with just the variable $r$ in our case and it is the violation percentage itself.

**5. Conclusion**

One of the area theories that attempt to address the sub-optimality decision making behaviour which recently emerged and received much recognition is the theories of limited attention or rational inattention. Most of these are founded upon axioms which make the validity of the predictions coming out of these theories depends upon the validity of the underlying axioms. We experimentally test the axioms underlying two of these new theories, those of Masatlioglu *et al.* (2012), and Lleras *et al.* (2017), which are based on the revealed preference framework. The experimental procedure elicits standard choice data. We observe the number of actual violations and then compare these with a ‘benchmark’ which was derived from simulations of random behaviour.

Out of the two weakenings of WARP, Lleras *et al*. is the more restricted version when compared to Masatlioglu *et al*. Masatlioglu *et al*. seems to perform the best in the consistencies analyses which is the main observation for the theory (axiom) violations. Lleras *et al*. performs poorly in term of direct violations. Particularly, it even performs worse than in the simulations of random behaviour. However, it did improve when full transitivity was assumed. WARP, which is the standard and normative way of describing choice behaviour, received some modest support from the data.

**Appendix A:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | *Px (x100)* | *X* | *Py (x100)* | *Y* | *E* |
| 1 | 50 | 19 | 50 | -3 | 8 |
| 2 | 80 | 6 | 20 | 17 | 8.2 |
| 3 | 60 | 18 | 40 | -3 | 9.6 |
| 4 | 70 | 8 | 30 | 13 | 9.5 |
| 5 | 10 | 19 | 90 | 7 | 8.2 |
| 6 | 20 | 16 | 80 | 7 | 8.8 |
| 7 | 30 | 14 | 70 | 7 | 9.1 |
| 8 | 40 | 11 | 60 | 8 | 9.2 |
| 9 | 90 | 7 | 10 | 17 | 8 |
| 10 | 80 | 13 | 20 | -3 | 9.8 |

**Appendix B:**

|  |  |
| --- | --- |
| Prob. No. |  Alternatives  |
| 1 | 1 |   |   |   |   |   |   |   |   |
| 2 | 2 |   |   |   |   |   |   |   |   |
| 3 | 3 |   |   |   |   |   |   |   |   |
| 4 | 4 |   |   |   |   |   |   |   |   |
| 5 | 5 |   |   |   |   |   |   |   |   |
| 6 | 6 |   |   |   |   |   |   |   |   |
| 7 | 7 |   |   |   |   |   |   |   |   |
| 8 | 8 |   |   |   |   |   |   |   |   |
| 9 | 9 |   |   |   |   |   |   |   |   |
| 10 | 10 |   |   |   |   |   |   |   |   |
| 11 | 1 | 2 |   |   |   |   |   |   |   |
| 12 | 3 | 9 |   |   |   |   |   |   |   |
| 13 | 4 | 9 |   |   |   |   |   |   |   |
| 14 | 5 | 9 |   |   |   |   |   |   |   |
| 15 | 5 | 10 |   |   |   |   |   |   |   |
| 16 | 1 | 2 | 3 |   |   |   |   |   |   |
| 17 | 1 | 2 | 6 |   |   |   |   |   |   |
| 18 | 1 | 2 | 9 |   |   |   |   |   |   |
| 19 | 1 | 2 | 10 |   |   |   |   |   |   |
| 20 | 1 | 3 | 9 |   |   |   |   |   |   |
| 21 | 1 | 5 | 9 |   |   |   |   |   |   |
| 22 | 3 | 5 | 10 |   |   |   |   |   |   |
| 23 | 3 | 9 | 10 |   |   |   |   |   |   |
| 24 | 4 | 5 | 10 |   |   |   |   |   |   |
| 25 | 4 | 6 | 9 |   |   |   |   |   |   |
| 26 | 4 | 8 | 9 |   |   |   |   |   |   |
| 27 | 5 | 6 | 9 |   |   |   |   |   |   |
| 28 | 5 | 8 | 9 |   |   |   |   |   |   |
| 29 | 5 | 8 | 10 |   |   |   |   |   |   |
| 30 | 1 | 2 | 3 | 10 |   |   |   |   |   |
| 31 | 1 | 2 | 4 | 9 |   |   |   |   |   |
| 32 | 1 | 2 | 5 | 6 |   |   |   |   |   |
| 33 | 1 | 3 | 5 | 9 |   |   |   |   |   |
| 34 | 1 | 3 | 5 | 10 |   |   |   |   |   |
| 35 | 1 | 3 | 9 | 10 |   |   |   |   |   |
| 36 | 1 | 5 | 6 | 9 |   |   |   |   |   |
| 37 | 1 | 5 | 7 | 9 |   |   |   |   |   |
| 38 | 2 | 3 | 9 | 10 |   |   |   |   |   |
| 39 | 2 | 4 | 5 | 10 |   |   |   |   |   |
| 40 | 2 | 4 | 6 | 9 |   |   |   |   |   |
| 41 | 3 | 4 | 5 | 10 |   |   |   |   |   |
| 42 | 3 | 4 | 6 | 9 |   |   |   |   |   |
| 43 | 3 | 5 | 6 | 9 |   |   |   |   |   |
| 44 | 3 | 5 | 6 | 10 |   |   |   |   |   |
| 45 | 3 | 5 | 9 | 10 |   |   |   |   |   |
| 46 | 4 | 5 | 8 | 9 |   |   |   |   |   |
| 47 | 4 | 5 | 8 | 10 |   |   |   |   |   |
| 48 | 4 | 8 | 9 | 10 |   |   |   |   |   |
| 49 | 1 | 2 | 3 | 5 | 10 |   |   |   |   |
| 50 | 1 | 2 | 3 | 9 | 10 |   |   |   |   |
| 51 | 1 | 2 | 4 | 7 | 9 |   |   |   |   |
| 52 | 1 | 2 | 4 | 8 | 9 |   |   |   |   |
| 53 | 1 | 2 | 5 | 6 | 8 |   |   |   |   |
| 54 | 1 | 3 | 4 | 5 | 10 |   |   |   |   |
| 55 | 1 | 3 | 5 | 6 | 10 |   |   |   |   |
| 56 | 1 | 3 | 5 | 7 | 10 |   |   |   |   |
| 57 | 1 | 3 | 5 | 8 | 10 |   |   |   |   |
| 58 | 1 | 3 | 7 | 9 | 10 |   |   |   |   |
| 59 | 1 | 3 | 8 | 9 | 10 |   |   |   |   |
| 60 | 1 | 5 | 6 | 7 | 9 |   |   |   |   |
| 61 | 2 | 3 | 4 | 9 | 10 |   |   |   |   |
| 62 | 2 | 3 | 5 | 6 | 9 |   |   |   |   |
| 63 | 2 | 3 | 6 | 9 | 10 |   |   |   |   |
| 64 | 2 | 3 | 7 | 9 | 10 |   |   |   |   |
| 65 | 2 | 4 | 5 | 6 | 10 |   |   |   |   |
| 66 | 3 | 4 | 5 | 7 | 10 |   |   |   |   |
| 67 | 3 | 4 | 6 | 9 | 10 |   |   |   |   |
| 68 | 3 | 5 | 6 | 7 | 10 |   |   |   |   |
| 69 | 3 | 5 | 6 | 8 | 10 |   |   |   |   |
| 70 | 3 | 5 | 7 | 9 | 10 |   |   |   |   |
| 71 | 4 | 5 | 7 | 8 | 9 |   |   |   |   |
| 72 | 4 | 6 | 8 | 9 | 10 |   |   |   |   |
| 73 | 1 | 2 | 3 | 4 | 8 | 9 |   |   |   |
| 74 | 1 | 2 | 3 | 4 | 9 | 10 |   |   |   |
| 75 | 1 | 2 | 3 | 5 | 6 | 10 |   |   |   |
| 76 | 1 | 2 | 4 | 7 | 9 | 10 |   |   |   |
| 77 | 1 | 2 | 5 | 6 | 7 | 8 |   |   |   |
| 78 | 1 | 2 | 5 | 6 | 8 | 10 |   |   |   |
| 79 | 1 | 3 | 4 | 5 | 6 | 10 |   |   |   |
| 80 | 1 | 3 | 4 | 5 | 8 | 10 |   |   |   |
| 81 | 1 | 3 | 4 | 5 | 9 | 10 |   |   |   |
| 82 | 1 | 3 | 4 | 6 | 9 | 10 |   |   |   |
| 83 | 1 | 3 | 5 | 6 | 8 | 10 |   |   |   |
| 84 | 1 | 3 | 5 | 6 | 9 | 10 |   |   |   |
| 85 | 1 | 4 | 5 | 6 | 7 | 9 |   |   |   |
| 86 | 1 | 5 | 6 | 7 | 9 | 10 |   |   |   |
| 87 | 2 | 3 | 4 | 5 | 6 | 9 |   |   |   |
| 88 | 2 | 3 | 4 | 5 | 7 | 10 |   |   |   |
| 89 | 2 | 3 | 4 | 5 | 9 | 10 |   |   |   |
| 90 | 2 | 3 | 4 | 8 | 9 | 10 |   |   |   |
| 91 | 2 | 4 | 5 | 6 | 7 | 10 |   |   |   |
| 92 | 2 | 4 | 5 | 6 | 8 | 10 |   |   |   |
| 93 | 3 | 4 | 5 | 7 | 9 | 10 |   |   |   |
| 94 | 1 | 2 | 3 | 4 | 5 | 7 | 10 |   |   |
| 95 | 1 | 2 | 3 | 4 | 5 | 8 | 9 |   |   |
| 96 | 1 | 2 | 4 | 5 | 6 | 7 | 9 |   |   |
| 97 | 1 | 2 | 4 | 5 | 6 | 7 | 10 |   |   |
| 98 | 1 | 2 | 4 | 5 | 7 | 9 | 10 |   |   |
| 99 | 1 | 2 | 5 | 6 | 7 | 9 | 10 |   |   |
| 100 | 1 | 3 | 4 | 5 | 6 | 7 | 9 |   |   |
| 101 | 1 | 3 | 4 | 5 | 6 | 7 | 10 |   |   |
| 102 | 1 | 3 | 4 | 5 | 6 | 8 | 10 |   |   |
| 103 | 1 | 3 | 4 | 5 | 6 | 9 | 10 |   |   |
| 104 | 1 | 3 | 4 | 5 | 7 | 9 | 10 |   |   |
| 105 | 1 | 3 | 4 | 5 | 8 | 9 | 10 |   |   |
| 106 | 1 | 5 | 6 | 7 | 8 | 9 | 10 |   |   |
| 107 | 2 | 3 | 4 | 5 | 6 | 8 | 9 |   |   |
| 108 | 2 | 3 | 4 | 5 | 7 | 9 | 10 |   |   |
| 109 | 3 | 4 | 5 | 6 | 7 | 9 | 10 |   |   |
| 110 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 10 |   |
| 111 | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 9 |   |
| 112 | 1 | 2 | 3 | 4 | 5 | 6 | 9 | 10 |   |
| 113 | 1 | 2 | 3 | 5 | 6 | 7 | 9 | 10 |   |
| 114 | 1 | 3 | 4 | 5 | 6 | 8 | 9 | 10 |   |
| 115 | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 10 |   |
| 116 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 117 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 |
| 118 | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

**Appendix C:**



**Instructions**

Preamble

Welcome to this experiment. Thank you for coming. Please read carefully these instructions. They are to help you to understand what you will be asked to do and how will you get paid. The experiment is simple and gives you the chance to earn a considerable amount of money. You will be paid in cash immediately after the experiment is completed.

The Experiment

The experiment is interested in how you take decisions. There are no right or wrong answers. You will be presented with a series of 118 problems, all of the same type. In each problem, there is a set of *lotteries*. We will describe in detail what we mean about a lottery in the next section. Your task is to choose one of these lotteries or not to choose any lottery at all in a problem. The outcome of playing out this lottery will lead to a *payoff* to you. Your payment for participating in this experiment will be the payoff from a randomly chosen one of these problems, (playing out the lottery of your choice), plus a £3 show-up fee. If it occurs that you did not choose any lottery in the randomly selected problem, your *payoff* will be your show-up fee. Details of all the payment procedures will be explained in the *payment* section.

A Lottery

We describe now what we mean by a ‘lottery’. Here we represent each lottery visually. The visual representation will be like the two examples below,





It is simplest to explain these in terms of the implications for your payment if one of these is randomly selected to be played out at the end of the experiment. What we will do in all cases is to ask you to draw ‒ without looking ‒ a disk out of a bag containing 10 disks numbered from 10, and an increments of 10, to 100. (You will be able to check that the bag contains all these disks before you do the drawing.) The number on the disk that you draw will determine a point on the horizontal axis; your payment would be the amount on the vertical axis implied by that point through the figure. At the point on the horizontal axis where the vertical axis changes it value, the payment would equal to the value of the vertical axis to the *left* of that point. In each lottery, there are two possible outcomes or *payoffs*.

So, for example, in the *top* lottery, if the number on the disk that you draw is between 10 and 50 *inclusive* you would get £19, notice that if the number on the disk is 50 you would get £19; if it is between 60 and 100 *inclusive* you would make a loss of £3. This loss will be deducted from your show-up fee. This implies that the chance of you getting paid £19 is 50 percent and the chance of you making a loss of £3 is also 50 percent. This will also be written in words. The caption will appear when you move the mouse cursor over the shaded areas. If the *bottom* lottery is to be played out, if the number on the disk that you draw is between 10 and 80 *inclusive* you would get £7, notice that if the number on the disk is 80 you would get £7; if it is between 90 and 100 *inclusive* you would get £11.

Let us give specific examples. In the *top* lottery, suppose the number on the disk that you draw is 70, then you would make a loss of £3 out of your show-up fee. In the *bottom* lottery, suppose the number on the disk that you draw is 30, you would receive £7.

Choices

In each problem, there is a set of *lotteries*. The number of lotteries varies from problem to problem. Your task is to choose one of these lotteries, or not to choose any lottery. You can choose a lottery by clicking at the box below the lottery of your choice. If you do not want to choose any lottery, you can do that by clicking the ‘*Prefer not to choose*’ button at the bottom part of the screen. Below is an example of a problem screen.



Payment

When you complete all 118 problems, please raise your hand and the experimenter will come to you. You will be lead to a separate room where the payment will take place. You will randomly choose one of the problems to play out for real. This is done by you drawing a disk from a bag containing 118 disks, each labelled number 1 to 118. The number on the disk that you draw is the problem that will be played out for real.

*If you chose one of the lotteries in that problem*

Your payment from the experiment will be from playing out a lottery of your choice from the randomly-chosen problem of the experiment plus the show-up fee of £3. You will randomly choose one numbered disk from another bag containing 10 disks numbered from 10, 20, 30, …, 100, and the number on the disk chosen will determine your payoff according to the procedure describe in the *lottery* section. If the *payoff* in the randomly chosen problem is zero you will receive only a show-up fee. If the *payoff* in the randomly chosen problem is negative, this will be deducted from your show-up fee. The maximum loss from a problem is -£3, therefore, at worst; you will be receiving £0 from this experiment.

*If you did not choose any lottery in that problem*

Your payment from the experiment will be only the show-up fee of £3.

What to do next (About the Experimental Software)

When you finish reading these Instructions, you should click on the ‘*start*’ button at the bottom of the screen (you will not be able to click this button until at least 5 minutes have passed). This will lead you to the actual experimental problems, and you will then be starting the experiment proper. Each problem screen has a countdown timer at the top right corner of the screen. You cannot click any button until 10 seconds have passed from when you started on that problem. There is a time limit of 45 seconds to make a decision on any problem. You can change your decision as many times as you want during this time period. You can click ‘*Submit*’ button before the time limit is reached. If you chose a choice and the time limit is over, that choice will automatically be your choice. If you do not choose any choices and the time limit is over, the default option, which is ‘*Prefer not to choose*’, will be taken as your choice on that particular problem.

*If you have any questions at any stage of the experiment, please raise your hand and an experimenter will come to you.*

*Thank you for your participation.*

Nuttaporn Rochanahastin

October 2017

Reference

Bouacida, E. and Martin, D. (2017) 'Predictive Power in Behavioral Welfare Economics'.

Caplin, A. and Dean, M. (2011) 'Search, choice, and revealed preference', *Theoretical Economics*, 6(1), 19-48.

Caplin, A. and Dean, M. (2015) 'Revealed Preference, Rational Inattention, and Costly Information Acquisition', *American Economic Review*, 105(7), 2183-2203.

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

Cherepanov, V., Feddersen, T. and Sandroni, A. (2013) 'Rationalization', *Theoretical Economics*, 8(3), 775-800.

Chetty, R., Looney, A. and Kroft, K. (2009) 'Salience and taxation: Theory and evidence', *American Economic Review*, 99(4), 1145-77.

De los Santos, B., Hortaçsu, A. and Wildenbeest, M. R. (2012) 'Testing models of consumer search using data on web browsing and purchasing behavior', *American Economic Review*, 102(6), 2955-80.

Lleras, J. S., Masatlioglu, Y., Nakajima, D. and Ozbay, E. Y. (2017) 'When more is less: Limited consideration', *Journal of Economic Theory*, 170, 70-85.

Manzini, P. and Mariotti, M. (2007) 'Sequentially rationalizable choice', *American Economic Review*, 97(5), 1824-1839.

Manzini, P. and Mariotti, M. (2010) 'Revealed preferences and boundedly rational choice procedures: an experiment', *Unpublished Paper*.

Manzini, P. and Mariotti, M. (2012) 'CATEGORIZE THEN CHOOSE: BOUNDEDLY RATIONAL CHOICE AND WELFARE', *Journal of the European Economic Association*, 10(5), 1141-1165.

Manzini, P. and Mariotti, M. (2014) 'Stochastic Choice and Consideration Sets', *Econometrica*, 82(3), 1153-1176.

Manzini, P., Mariotti, M. and Tyson, C. J. (2013) 'Two‐stage threshold representations', *Theoretical Economics*, 8(3), 875-882.

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

Matějka, F. and McKay, A. (2015) 'Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model', *American Economic Review*, 105(1), 272-98.

Selten, R. (1991) 'Properties of a measure of predictive success', *Mathematical Social Sciences*, 21(2), 153-167.

Simon, H. A. (1955) 'A Behavioral Model of Rational Choice', *The Quarterly Journal of Economics*, 69(1), 99-118.

Sims, C. A. (1998) 'Stickiness', in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, 317-356.

Sims, C. A. (2003) 'Implications of rational inattention', *Journal of monetary Economics*, 50(3), 665-690.

Sims, C. A. (2010) 'Rational inattention and monetary economics' in *Handbook of Monetary Economics*Elsevier, 155-181.

Tyson, C. J. (2008) 'Cognitive constraints, contraction consistency, and the satisficing criterion', *Journal of Economic Theory*, 138(1), 51-70.

Tyson, C. J. (2013) 'Behavioral implications of shortlisting procedures', *Social Choice and Welfare*, 41(4), 941-963.

Tyson, C. J. (2015) 'Satisficing behavior with a secondary criterion', *Social Choice and Welfare*, 44(3), 639-661.

1. Caplin (2016) provides a useful and comprehensive review. [↑](#footnote-ref-1)
2. The property of the attention filters is that the consideration sets are unaffected when an alternative the DM does not pay attention to becomes unavailable. [↑](#footnote-ref-2)
3. If an alternative was considered in a set then it must be considered in its subset. [↑](#footnote-ref-3)
4. Consideration set formation is crucial in rational inattention but it is very difficult to observe or pin down. [↑](#footnote-ref-4)
5. The design in this paper focuses only on choice data. [↑](#footnote-ref-5)
6. Lotteries details can be found in Appendix A. [↑](#footnote-ref-6)
7. Randomised lotteries in each problem can be found in Appendix B. [↑](#footnote-ref-7)
8. for example, Grether (1978), Grether and Plott (1979). [↑](#footnote-ref-8)
9. Note that these are all *direct* inferences. We do not assume transitivity at this point. [↑](#footnote-ref-9)
10. Note that we use square brackets to denote this type of relationship. [↑](#footnote-ref-10)
11. This refers to the cycle of length two as transitivity is still not assumed at this point. [↑](#footnote-ref-11)
12. Note that we use round brackets to denote this type of relationship. [↑](#footnote-ref-12)
13. The extraction procedures are similar to example provided in table 2. [↑](#footnote-ref-13)
14. Percentage points. [↑](#footnote-ref-14)
15. And among repetitions in case of simulation. [↑](#footnote-ref-15)
16. The procedures are similar to those examples in table 3. [↑](#footnote-ref-16)
17. The procedures are similar to those in table 5. [↑](#footnote-ref-17)