

Gaudeul, Alexia; Crosetto, Paolo

**Working Paper**

## Fast then slow: A choice process explanation for the attraction effect

cege Discussion Papers, No. 386

**Provided in Cooperation with:**

Georg August University of Göttingen, cege - Center for European, Governance and Economic Development Research

*Suggested Citation:* Gaudeul, Alexia; Crosetto, Paolo (2019) : Fast then slow: A choice process explanation for the attraction effect, cege Discussion Papers, No. 386, University of Göttingen, Center for European, Governance and Economic Development Research (cege), Göttingen

This Version is available at:

<https://hdl.handle.net/10419/209649>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*


**FAST *THEN* SLOW: A CHOICE PROCESS  
EXPLANATION FOR THE ATTRACTION  
EFFECT**

---

Alexia Gaudeul<sup>Ⓒ</sup>  
Paolo Crosetto

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

# Fast *then* Slow: A Choice Process Explanation for the Attraction Effect\*

Alexia Gaudeul<sup>†</sup>  Paolo Crosetto<sup>‡§</sup>

December 17, 2019

## Abstract

In this paper we provide choice-process experimental evidence that the attraction effect is a short-term phenomenon, that disappears when individuals are given time and incentives to revise their choices. The attraction (or decoy) effect is the most prominent example of context effects, and it appears when adding a dominated option to a choice set increases the choice share of the now dominant option at the expense of other options. While widely replicated, the attraction effect is usually tested in hypothetical or payoff-irrelevant situations and without following the choice process. We run a laboratory experiment where we incentivize choice, vary the difference in utility between options and track which option participants consider best over time. We find that the effect is a transitory phenomenon that emerges only in the early stages of the choice process to later disappear. Participants are fast *then* slow: they first choose the dominant option to avoid the dominated decoy and then progressively revise their choices until choice shares come to correspond to price differences only. We expand our analysis by considering differences in utility among options and differences in the presentation of options (numerical or graphical). We also consider differences in the choice processes followed by individuals (intuitive vs. deliberative). This allows us to ascribe more precisely the role of fast and slow cognitive process in the emergence and disappearance of the attraction effect.

**Keywords:** asymmetric dominance, attraction effect, induced preferences, choice process, time constraint, rationality, context effects

**JEL:** C91; D12; D83

---

\*This paper was presented at the Annual Conference of the Vereins für Socialpolitik in Leipzig, 2019, at the 12th Maastricht Behavioral and Experimental Economics Symposium (M-BEES) in Maastricht, 2019, at the Sixth International Meeting on Experimental and Behavioral Social Sciences (IMEBESS) in Utrecht, 2019, at the Workshop on Choice Process Data of the ESA North American Meeting in Antigua, Guatemala, 2018, at the ESA World Meeting in Berlin, 2018, at the Tagung der Sozialwissenschaftlichen Ausschusses des Vereins für Socialpolitik in Essen, 2018 and at the GfW-Jahrestagung in Kassel, 2017. It was also presented at invited seminars at the Université Paris 8, Saint-Denis, France, GREDEG-CNRS, Université of Nice Sophia Antipolis, France, Universität Duisburg-Essen, Germany, and the Department of Economics and Management of the University of Pisa, Italy.

<sup>†</sup>Faculty of Economic Sciences, Georg-August-Universität Göttingen, 37073 Göttingen, Germany. email: a.gaudeul@gmail.com

<sup>‡</sup>Univ. Grenoble Alpes, INRAE, CNRS, Grenoble INP, GAEL, 38000 Grenoble, France. email: paolo.crosetto@inra.fr

<sup>§</sup>The order of authors is certified random, see [Ray and Robson \(2018\)](#).

# 1 Introduction

Since Tversky (1972) we know that adding irrelevant alternatives to the consideration set impacts choices – a phenomenon known as *context effects*. Context effects are direct violations of the axiom of Independence of Irrelevant Alternatives (“IIA”), and as such constitute a powerful challenge to rational choice theory. They also have crucial practical applications in marketing and consumer protection, since they show that choices can be influenced by careful engineering of the choice set. A particularly prominent example among several documented context effects (Tversky and Simonson, 1993) is the *attraction effect*, also known as *decoy* or *asymmetric dominance effect* (henceforth “AE”, Huber et al. (1982)). Under the AE, adding an option that is clearly dominated (*decoy*) by one option in a consideration set triggers an increase in the choice share of the now dominant option (*target*) at the expense of the other option (*competitor*) – even though the decoy should indeed be irrelevant.

Due to its theoretical importance as a counter-example to rational choice theory, its counter-intuitive nature, and its obvious potential for marketing practices, the AE has generated an enormous literature over the thirty-five years of its existence. The AE has been widely replicated in marketing and consumer research (Huber and Puto, 1983; Simonson, 1989; Park and Kim, 2005; Malkoc et al., 2013), cognitive psychology (Trueblood et al., 2013), neuroscience (Hu and Yu, 2014), game theory (Wang et al., 2018), experimental economics (Herne, 1999; Sonsino, 2010; Kroll and Vogt, 2012; Castillo; Sürücü et al., 2019), and even in some studies on animal behavior (Schuck-Paim et al., 2004; Shafir et al., 2002).

Despite the large amount of evidence, the AE has shown little robustness – to the point that its applicability as a marketing tool in the real world and its theoretical importance have been put into question. The effect is muted when individuals have information about brands or are in other ways familiar with the products (Ratneshwar et al., 1987) or when the product description is unambiguous and precise (Mishra et al., 1993). It disappears or is severely reduced when the options are presented graphically rather than numerically (Frederick et al., 2014; Yang and Lynn, 2014), when the products have negative rather than positive attributes (Malkoc et al., 2013; Chang et al., 2015), when individuals are not indifferent among options before the introduction of a decoy (Crosetto and Gaudeul, 2016; Farmer et al., 2016), and, among animals, when marginally changing the usual design (Cohen and Santos, 2017). The AE also failed to replicate in a recent large experiment with real-world consumer choices (Trendl et al., 2018) that complied with all the conditions laid down by the original authors to allow successful replication (Huber et al., 2014). On the other hand, the effect is amplified when individuals are asked to justify their choices (Simonson, 1989) and when the dominance relation is more focal (Mishra et al., 1993; Król and Król, 2019).

Several theories of choice predict an effect like the AE in the presence of dominance. The effect can be explained by reference-dependent utility coupled with loss aversion (Usher and McClelland, 2004), decision field theory (Roe et al., 2001), salience theory (Bordalo et al., 2013), using the elimination by aspect theory of Tversky (2003), assuming shifting decision weight across attributes (Ariely and Wallsten, 1995), or assuming a form of trade-off aversion (Hedgcock and Rao, 2009). However most of those theories cannot account for the limited robustness of the AE nor explain which factors reduce or mute it.

A different approach is provided by thinking of the AE as the result of a heuristic or a short-term, intuitive decision rule. Heuristics are simple choice rules that reduce the complexity of a problem by ignoring most information yet enable fast decision and lead to sufficiently good results (Gigerenzer et al., 2000). Heuristics are associated with the “satisficing” (Simon, 1959), “intuitive”, “fast” (Kahneman, 2011), or “compensatory” (Tversky, 1972; Dieckmann et al., 2009) thinking modes of dual-mode decision theories. Under this view, the AE, despite being a violation of rational choice could be locally optimal, especially in cases in which there is little information, the signals are noisy, or the cognitive

abilities of the decision maker do not allow for sufficient accuracy.

In this paper we ask whether the attraction effect is a short-term phenomenon, that disappears when individuals are given time and incentives to revise their choices, or is a persistent feature of choice even in the long run and in the face of incentives. Answering this question is relevant both for the literature on context effects and the economic literature at large, as it provides insights about the lack of robustness of the AE and helps in assessing the seriousness of the threat posed by context effects to rational choice theory.

The existing literature is ill-equipped to answer this question. First, in most existing papers choices are not incentivized, giving little reason for participants to revise an intuitive first response. According to [Lichters et al. \(2015\)](#), out of 52 reviewed studies only one was incentivized. Second, *all* replications of the AE except [Crosetto and Gaudeul \(2016\)](#) and [Farmer et al. \(2016\)](#) rely on choices in which the target and the competitor are carefully chosen to make most people indifferent. Because a person with strong preference for one item would be less affected by dominance relations, virtually all the existing literature tries to ensure indifference between options, in which case any cue might impact choice. Third, the large majority of studies exploits between-subjects designs, thus allowing researchers to investigate the effect only at the aggregate level, hiding individual-level differences and barring all insights on individual decision processes and switches between thinking modes. Finally, only two papers impose time pressure ([Dhar et al., 2000](#); [Pettibone, 2012](#)). Their results indicate that context effects (the AE for [Pettibone \(2012\)](#), the compromise effect for [Dhar et al. \(2000\)](#)) are reduced under time pressure, thus hinting at the fact that context effects might be persistent and long-term features of decision processes. Nonetheless, both papers impose an exogenous, fixed time for participants to submit their choices, and analyze between-subjects aggregate data. The issue there is that aggregate results could be an artifact of unobserved subject heterogeneity in response time, cognitive ability, thinking styles and accuracy, since the exogenously imposed time limit could generate no time pressure on some participants while at the same time exerting too much pressure on others.

In this paper we introduce a novel incentivized experimental design that allows us to measure the AE at the individual level against an optimal benchmark and to observe, for each subject and with endogenous time pressure, the switch from “fast” to “slow” thinking modes in conditions of indifference — to replicate existing findings — but also away from indifference. We provide strong choice-process experimental evidence that supports the view that the AE can best be understood as an intuitive strategy that relies on dominance to provide a first, approximate, fast answer, but that is likely to be abandoned upon closer inspection and given time and incentives.

We use a state-of-the art choice process elicitation mechanism ([Caplin et al., 2011](#); [Spiliopoulos and Ortmann, 2017](#)) to observe not only the final choices, but a continuum of choices from the first time a subject sees a menu of options to the last time he can make or revise his choice. Subjects are incentivized to give, at any moment in time, what they think is the best option at that time; they are also free to revise their choice at any later moment. We induce preferences and incentivize choices. This allows us to avoid hypothetical bias, to explore the robustness of the AE away from indifference, and to set up an objective standard for each choice. In our task, given enough time, effort and accuracy, participants can objectively find the (unique) optimal option. We hence have a task in which the use of System 2, slow, deliberative decision processes can lead to the optimal solution eventually. This allows us to track the transition from intuitive to reasoned thinking – the strategy used by participants to navigate the speed-accuracy trade-off ([Wickelgren, 1977](#)) — in a task in which both modes are possible. Finally, we expose participants to several dozens tasks, so that we can estimate the choice process of each subject separately and tell apart types who follow different choice processes as well as their relative shares in our sample.

Our results support the idea that the attraction effect is best viewed as a short-term phenomenon.

First, on the aggregate data, we show that the AE appears within the first few seconds of exposure to a problem, when participants submit fast, intuitive choices, and later declines as participants use their time to revise their decision. Second, at the individual level, we show that this rise-and-fall pattern is driven by the interaction of three classes of participants: *heuristic* types, that use dominance and then stop searching; *maximising* types, who rely only on their ability to make accurate price comparison without appearing to exploit dominance; and *fast then slow* types, that exploit dominance first to then go on to rely on more informed decision later. Third, we confirm that the AE is strongest in situations of indifference between the target and the competitor – indeed, this is the only situation in which a (small) AE persists after the initial seconds and is still present at the end of the allotted time. Fourth, we cannot replicate in our settings the finding of [Frederick et al. \(2014\)](#), whereby participants would be less prone to the AE in presence of graphical rather than numeric stimuli.

Our finding about the short-term nature of the AE can provide answers to some unresolved questions in the literature about its relevance to both theory and practical applications.

In theoretical terms, thinking of the AE as a short-term decision strategy limits its role in reshaping existing theories of rational choice; as most heuristics, it is locally rational, and if applied in its proper context and in presence of bounded rationality, limited accuracy or noisy information it improves pay-offs; as most heuristics, it does not invalidate more deductive and comprehensive accounts of rational choice theories outside of its limited scope. The IIA axiom is alive and well, given enough time and incentives. After a few seconds, all but a minority share of heuristic-only decision makers switch to more reflective and analytic decision modes, thus abandoning the AE-driven choice and exploiting their ability to make trade-offs to make better informed choices.

In terms of the practical applicability of AE research, our results imply that the AE is bound to have a small to negligible effect on choice – bar on the few situations in which dominance really stands out. First, the AE faces competition from other heuristics. In most applied contexts, it is likely to be washed out by other cues, such as familiarity, brand preferences, habits, and other cues. Second, in most contexts it is hard to put the AE to use. For the AE to work, the dominance relation must be unambiguous and focal. In most experiments the dominance relationship is artificially enhanced to be central, clear, and unmissable. This is likely not to be the case in applied settings.

## 2 Hypotheses

Our main aim is to better define the theoretical nature of the AE, either as a short-term, intuitive strategy or as a persistent feature of choice. If the AE is a stable feature of choice, then it should appear both when participants are given a short or a long time to submit their choices, and whether participants are incentivized or not. Indeed, it should be more apparent the longer the participants are given time and incentives to apply effort to the problem. This is the message of the existing literature on the impact of time pressure on choice in menus with dominated alternatives ([Pettibone, 2012](#); [Dhar et al., 2000](#)). According to that literature, the attraction effect takes time to become established as participants compare options. On the other hand, if the AE is a transitory phenomenon based on a fast and frugal heuristic, we would expect it to be prevalent in the first stages of the choice process, to then be superseded by other, considered and slow choice strategies. Our main hypothesis, focusing on the dynamic time path of the AE, is then that

**Hypothesis 1.** *The AE will be more pronounced in the early stages of the decision process, to then decrease over time.*

The relevance of the AE depends on several factors. First, the AE needs dominance to be clear and focal, and should thus be more present in the absence of graphical stimuli that might give other,

alternatives cues to guide choice. This is in line with the findings of [Frederick et al. \(2014\)](#) and [Yang and Lynn \(2014\)](#) that show little effect of the AE with graphical stimuli. We hence run a graphical and a numeric treatment, and hypothesize that the AE will be lower in the graphical treatment, where dominance is harder to assess and stimuli are fuzzier:

**Hypothesis 2.** *The AE will be more pronounced when options are presented numerically rather than graphically.*

Second, the AE has been shown to be more pronounced in conditions of indifference, and to decrease when moving away from it ([Farmer et al., 2016](#); [Crosetto and Gaudeul, 2016](#)). This might be due to two complementary reasons. First, the dominance heuristic may be crowded out by the search for the lowest price as the payoff distance between the target and the competitor increases. Second, participants may rely more on analytic decision modes and less on short-term decision strategies when there are actual differences in price across options. In both cases, we expect a reduction of the AE as the payoff distance between the options increases:

**Hypothesis 3.** *The AE will be more pronounced when the choice between the target and the competitor is indifferent, and less pronounced when the two options differ in price.*

Beyond those hypotheses, we also exploit the repeated individual observations to further investigate differences in decision modes across participants, and the determinants of the decision by participants to defer choice ([Dhar, 1996](#); [Dhar and Nowlis, 1999](#); [Anderson, 2003](#)) and not to rely on heuristics at all. We will hence show if the aggregate dynamic pattern of the AE hypothesized in 1 is due to differences in the speed of making a first choice between heuristic and optimizing participants, or if indeed some people *do* change decision modes and switch from heuristics to accurate estimation.

## 3 Experimental design

### 3.1 Expenditure minimization task

Participants in our experiment performed an intuitive *expenditure minimization task*. They were told they had to buy three liters of gasoline. They were given a budget of 5 € and kept as a payoff all the money not spent on the task after they had bought 3 liters of gasoline, choosing among the different offers on display. We did not provide the participants with prices per liter of gasoline. We instead split the basic information needed by the subject – price per liter – into two different dimensions: quantity and price per such quantity. By doing so, we opened up a two-dimensional continuum for each unit price, as, say, a price of 2 € per liter can be displayed as 1 € for half a liter, 3 € for 1.5 liters, and so on.

This design replicates most features of traditional AE designs, at the same time overcoming some of their limits.

Traditional AE designs propose a choice over items defined over two dimensions (location and size of apartments, quality and price of beers, resolution and durability of tv-sets) and participants must assess the utility trade-off of the two dimensions. Our task allows us to move the difficulty of combining multi-dimensional attributes from the (unobserved and not measurable) utility space to the cognitive difficulty of making price/size evaluations over (measurable) money. The presence of money allows us to evacuate preferences and objectively measure performance. The task is indeed mono-dimensional – participants care about one dimension only, money – but the cognitive difficulty of comparing different sizes and prices makes it two-dimensional, as long as the size/price evaluations involved are not trivial. The use of money allows us to move from the aggregate (between-subjects) to the individual (within-subjects) level, and to test the behavior of each subject individually for the presence of the AE; we hence directly test the IIA axiom at the individual level.

The between-participants nature of the traditional design implies that differences in preferences across groups could affect the results. This shortcoming is usually dealt with by carefully choosing the stimuli so that most participants are nearly indifferent between the target and the competitor. Our design relies on induced preferences. By letting the price of the different options vary, we can seamlessly change the incentives and observe the behavior of participants at the indifference point – as in the existing literature - but also away from it.

### 3.2 Treatments

We employ a mixed between- and within-subjects treatment structure.

Between subjects, we vary the stimuli used to visualize the expenditure minimization task. In the *graphical* treatment (Figure 1, left) the quantity of gasoline of each option was displayed graphically by means of a partially filled jerry-can, the part filled (in light red) indicated the quantity, and a dashed line indicated the target quantity (3 liters) to be bought. In the *numeric* treatment (Figure 1, right), the quantity was displayed as a simple number. This variation is inspired by the recent controversy [Frederick et al. \(2014\)](#); [Huber et al. \(2014\)](#); [Yang and Lynn \(2014\)](#) over the robustness of the AE when the options are presented in a graphical or verbal rather than numerical form, and allows us to test Hypothesis 2.

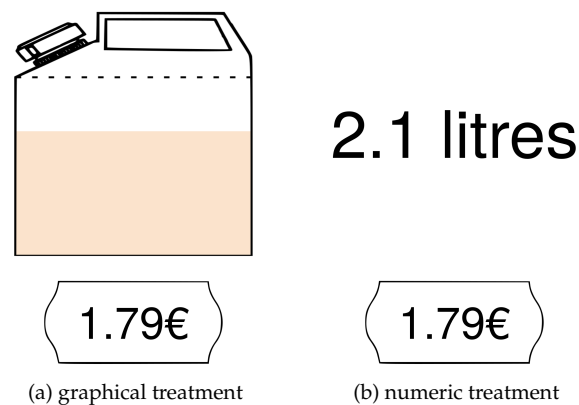


Figure 1: Stimuli used in the main task

Within subjects, we vary across repetitions of the base task the relative price of the target and competitor options. The competitor is allowed to have a price that ranges from 85% to 115% of the target price. This range includes the special case of indifference that is studied by the bulk of the literature, but allows us to study the AE also in situation in which it has monetary consequences. This within-subject variation allows us to observe, for each subject, the cost incurred for following the AE away from indifference, if any, and provides evidence to test Hypothesis 3.

Again within subjects, we provide two different measures of the AE. Traditionally, the AE is measured as the difference, between-subjects, in the choice share of the target between a situation in which participants choose among two options (target and competitor) and one in which they choose among three options (target, competitor and decoy). We replicate this measure at the individual level, exposing participants to both types of choices. We call this traditional measure *3vs2*. Despite the large literature supporting it, this measure relies on choice sets of different size, thus potentially introducing a bias. We hence add a second measure of the AE that is built on choice sets of the same size. This measure has been used recently in the rare within-subjects studies of the AE, for instance by [Trueblood et al. \(2013\)](#) and [Farmer et al. \(2016\)](#) – we call this measure *3vs3*. We measure the AE as the difference in the choice



share of the target between choice sets with three options with and without a decoy. We implement this by hiding the dominance relation between the target and the decoy – choice sets without a decoy are just choice sets with a bad option that is not clearly dominated by the target. An example of two choice sets used to compute the traditional *3vs2* AE measure is given in Figure 2, where the added central option is the decoy, that is dominated by the target (left) and in no dominance relation with the competitor (right). An example of two choice sets used to compute the new *3vs3* AE measure is given in Figure 3, where the decoy, in the center, is not clearly dominated in the upper choice set while it is clearly dominated by the left option in the lower choice set.

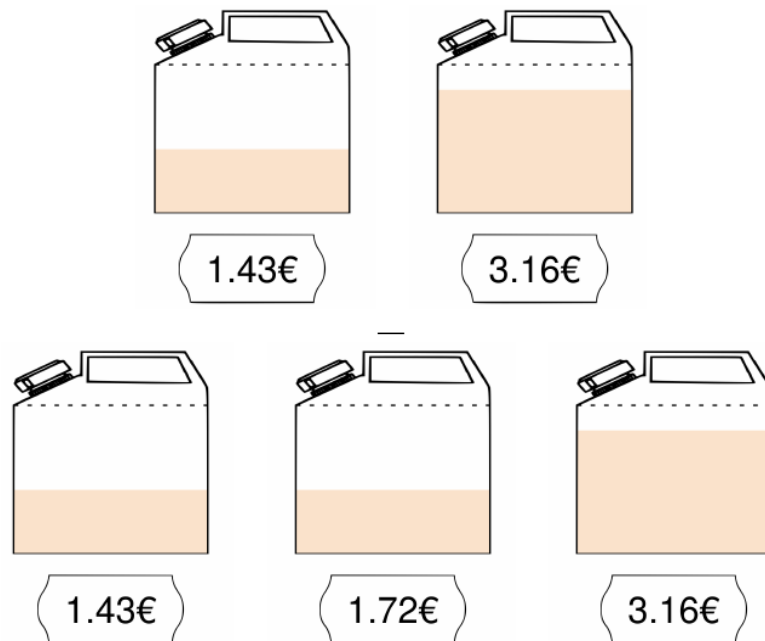


Figure 2: Example of choice sets used to obtain the *3vs2* AE measure

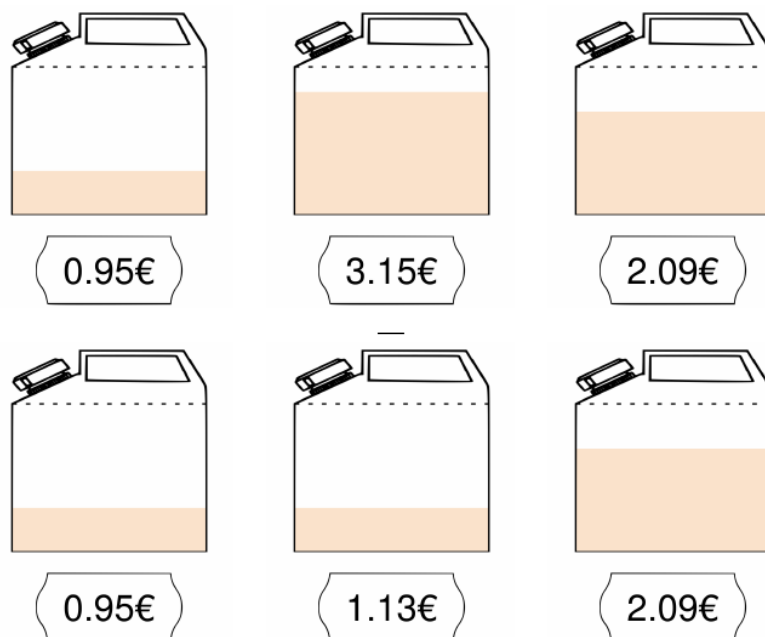


Figure 3: Example of choice sets used to obtain the *3vs3* AE measure

### 3.2.1 Details of the creation of offers and choice sets

Subjects were exposed to 40 expenditure minimization tasks. Across tasks we varied the relative price of the target and the competitor, the presence or absence of a decoy (3vs2) and the focality of the dominance relation between target and decoy (3vs3).

In our design each option is fully described by two attributes: the unit price of gasoline and the amount of gasoline sold. The price that was shown to the participants was then simply computed as unit price times amount. We started by creating 20 choice sets made up of three options as follows:

- We drew four different unit prices for the target from a uniform distribution in the open interval  $[0.6, 1.4]$ , with a granularity of 5 cents. This unit price was then multiplied by 1.2 to obtain the price of the decoy. That is, the decoy is *always* 20% more expensive than the target. The unit price of the competitor was derived from the unit price of the target, multiplied by 0.85, 0.95, 1, 1.05, or 1.15. The distribution of those unit prices around the price of the first option is hence symmetric, and varies from no (0%) to slight (5%) and strong (15%) negative or positive price differences. The competitor is therefore *always* cheaper than the decoy, but in varying degrees; and it might be cheaper or more expensive than the target. Overall, we thus obtain  $4 \times 5 = 20$  basic choice sets.
- For each of these 20 choice sets we then drew the amount of gasoline of the target, competitor and decoy from the set  $\{0.9, 1.3, 1.7, 2.1, 2.5\}$  (liters), without replacement. This guarantees that no option has the same size.
- When draws of unit price and amount of gasoline resulted in two offers in a choice set having identical shown prices (i.e., *unit price*  $\times$  *size*), then this choice set was discarded and the procedure repeated so as to avoid having price decoys – i.e. a dominance relation on the shown prices. Dominance relations could then only appear in the size dimension.

Each of those menus was transformed into a menu with three options and a decoy by simply assigning to the decoy the same size as the target and adjusting the shown price accordingly. We then numbered menus from 1 to 20. For even menus, a corresponding menu with two options was generated by simply removing the decoy from the choice set. For odd menus, the corresponding menu was the original menu with no decoy. The position of the options in the resulting 40 screen was randomized once, and then applied to all participants.

A detailed list of the 40 choice sets, in their graphical and numeric versions, as well as a collection of screenshots are available on the OSF page of this paper (<https://osf.io/xr28d/>).

### 3.3 Incentivized elicitation of the choice process

We not only recorded final choices, but also incentivized participants to reveal the option they consider best over time, by using a choice process elicitation method adapted from Caplin et al. (2011). Other methods to track the choice process, such as eye-tracking (Reutskaja et al., 2011) or mouse-tracking Lohse and Johnson (1996), allow one to see what decision-makers look at, but not what option they think best at each point of time. By incentivizing choice over time, we obtain information about what a participant would have chosen under different degrees of time pressure, without having to exogenously impose a time limit in order to elicit intuitive decisions.

Participants were allotted 20 seconds for each choice set, which were sufficient for even the slowest participants to make a sufficiently considered decision. Participants could change their choice at any point in time. For each second  $t$ , their choice  $c_{it}$  is recorded as their most recently chosen option.<sup>1</sup> At

<sup>1</sup>The participants therefore do not need to click every second on their most preferred option. They simply click when they want to make a choice, and when they want to change their choice.

the end of the allotted time, the data obtained from each subject is a vector containing all the choices,  $\mathbf{C}_i = \{c_{it} | t = 1 \dots T\}$ . One time point  $\bar{t}$  is then uniformly drawn,  $\bar{t} \sim U(1, T)$ , and the choice recorded at that time  $c_{i\bar{t}}$  is binding and determines the subject's payoff. If no choice had been submitted by time  $\bar{t}$ , then  $c_{i\bar{t}} = NA$ , and the participant is assigned a uniform random choice in the alternative space,  $c_{i\bar{t}} \sim U(\text{target}, \text{competitor}, \text{decoy})$ . This was made clear to participants in the instructions. To increase the participants' familiarity with this elicitation mechanism, the participants went through 4 non-incentivized tasks before the start of the experiment, and received feedback about their choices, the random drawing of the binding time point, and the resulting implemented choice.

This elicitation mechanisms incentivizes the participants to submit a choice as soon as they think to have improved on choosing at random. This is particularly relevant in the presence of decoys, as those are clearly dominated. A first, fast choice reduces the chance that the decoy will be chosen by the random mechanism. Once a first choice is submitted, participants are incentivized to reconsider and improve, if possible, on their choice; the earlier a participant settles on what he thinks is the best option, the higher the probability that this option will be the one actually implemented; nonetheless, the subject continuously faces an incentive to change his mind upon further reflection. With respect to a normal choice task, participants hence reveal their view of what is the optimal choice over time. Moreover, with respect to time pressure tasks, the time constraint is not exogenously set by the experimenter, but endogenously determined by each subject, thus allowing participants to be heterogeneous in their reaction, computation, and response times.

## 4 Experimental details and procedures

We ran seven experimental sessions involving a total of 111 participants, 63 for the graphical and 48 for the numeric treatment. The sessions took place at the GAEL experimental economics laboratory in Grenoble, France, in July 2017. Participants were recruited from the general population with ads in local newspapers as well as from an existing database of potential participants in and around Grenoble, a mid-sized French city with a metro area of about half a million people.

All sessions followed the same script. Upon entering, participants were randomly assigned a code and seated. Instructions were read aloud and displayed on the participants' computer screens. After instructions were read and all questions answered, participants went through 4 practice screens. The screens used different stimuli, but were otherwise identical to the ones used in the main task. Out of the 4 screens, two showed choice sets of three choices with a decoy, one a choice set of three choices with no decoy, and another a choice set of two choices with no decoy. At the end of the 4 practice tasks, subject saw a feedback screen, giving them information about the second randomly chosen to be binding, whether at that second they had or not submitted a choice, their choice at that moment (if no choice, the option randomly chosen by the computer), the total cost of the gasoline and their profit. After all remaining questions, if any, had been answered, participants moved to the main task.

In the main task, participants saw a blank screen with a time counter for four seconds. Then, the stimuli, with no possibility to choose, were shown for further 4 seconds. Then the screen became active and the time bar at the bottom of the screen started filling up. Participants faced the screen for 20 seconds; then the cycle started again. It took about 20 minutes to cycle through the 40 decision menus. The order of the 40 tasks was randomized across participants. The order of the options on the screen was also randomized, but fixed for all participants.

At the end of the main task participants were asked to fill in 5 different questionnaires. These were

- a socio-demographic questionnaire asking question about gender, education, income, profession;

- SOEP questions on general attitude to risk (Dohmen et al., 2011), trust Dohmen et al. (2008)) and a question measuring loss aversion;
- a qualitative questionnaire to evaluate participants’ understanding of the task and inquire into possible experimenter demand effects;
- a 3-item Cognitive Reflection Test (CRT) questionnaire (Frederick, 2005);
- a Consumer Confusion Proneness questionnaire (Walsh et al., 2007).

The English translation of the original French instructions is in appendix F. The web-based experimental software, written in php, as well as the original French instructions are available upon request.

Given that we recruited from the general population, sessions were run at convenient times for the participants, either during lunch breaks or in the evening. The distribution of the two treatments across times of the day was balanced.

Participants received a 10 € show-up fee. Moreover, five randomly picked out of the 40 menus were individually and independently drawn to be payoff relevant. Given this payoff rule, participants could earn up to a theoretical maximum of 15.49 € in addition to the show-up fee. This would happen if they always made the profit maximizing choice and the choice situations with the lowest prices were randomly drawn for payment. In practice, subject earned on average 10.00 € in addition to the show-up fee (st.dev. 1.57). Payoffs in the two treatments were virtually identical.

## 5 Results

In this section we present first aggregate results focusing on mean behavior and hence implicitly assuming that all participants follow the same patterns. Section 5.1 displays descriptive statistics of participants’ aggregate choice patterns in time for the two treatments and over the different within-subjects variations. To gain a deeper insight on behavior and being able to model jointly the choice made and the time patterns in which the choice was made, section 5.2 presents a structural model of choice that allows us to give a value to the ‘no choice’ option and to jointly estimate the AE alongside accuracy. The results of this model are presented in section 5.3.

In section 5.4 we show that the assumption of similar choice patterns across individuals does not hold in our sample. Subjects clearly show different choice strategies, different uses of the dominance heuristic, different accuracy and a different use of time. In section 5.4.1 we then propose and estimate a mixture model able to account for the different types present in our population.

### 5.1 Aggregate behavior

The main measure of the attraction effect used in the literature is the difference in the choice share of the *target* across choice sets with and without a decoy. That is, if the target is chosen 50% of the time when there is no decoy and 55% of the time when there is a decoy, then the AE is 5%. As explained page 3.2, we measure this in two ways: by comparing choices from a set with two items to the choices from a set with three items, one of which is dominated – this is the *3vs2* measure – and by comparing choices from a set with three items, none of which is dominated, to the choices from a set with three items, one of which is dominated – this is the *3vs3* measure. For both *3vs3* and *3vs2* we then add a time dimension: the difference is computed for each of the twenty seconds the participants spent on the task. Finally, we compute the difference separately for the *numeric* and *graphical* treatments.

Figure 4 presents the results of both the *3vs2* and *3vs3* measures (rows) for the numeric and graphical treatments (columns) over time. Lines represent the average difference in percentage points in the

choice share of the target between a situation in which a dominated decoy is present and one in which it is not; error bars represent 95% confidence intervals.

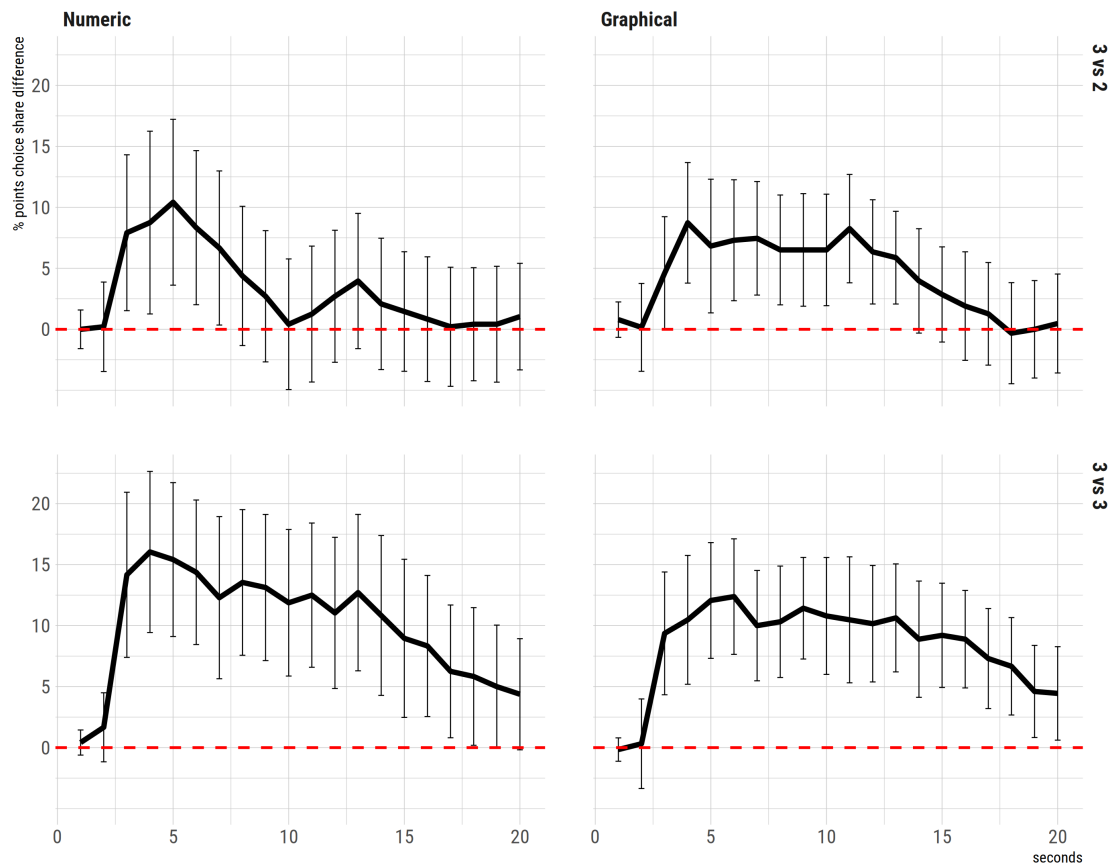


Figure 4: AE in time across different measures and treatments

We observe from those graphs that the AE is a transitory effect. This is true for both visual representations of our stimuli and for both measures. In all cases, the AE rises in the first seconds, when participants submit their first choices, and falls afterwards, when participants revise their choices and/or slower participants start choosing. The size and permanence of the attraction effect appears to be higher in the numeric treatment, and the AE appears to be higher when comparing 3vs3 menus than when comparing 3vs2 menus. We will test later the significance of those difference.

Figure 4 disregards the role of the relative value of the target and the decoy. The figure aggregates data from choice sets in which the relative price of the target and the competitor vary. This is justified because the distribution of the price of the competitor is symmetric around the situation of indifference, in which the real price of the target and the competitor are the same, and the measure shown is a difference between two identical situations, bar the presence of a dominated decoy, and hence every remaining possible bias is eliminated. Nonetheless, the likelihood of observing an attraction effect may depend on the relative price of the target with respect to the competitor.

Figure 5 therefore shows the dynamics of the AE for both measures at different levels of relative price of the target, from costing 15% more to 15% less than the competitor. While the lower number of observations means that confidence intervals are wider, we recognize the same general pattern of rise and fall of the attraction effect over time. The target enjoys a *transitory* increase in choice share at every level of its relative price.

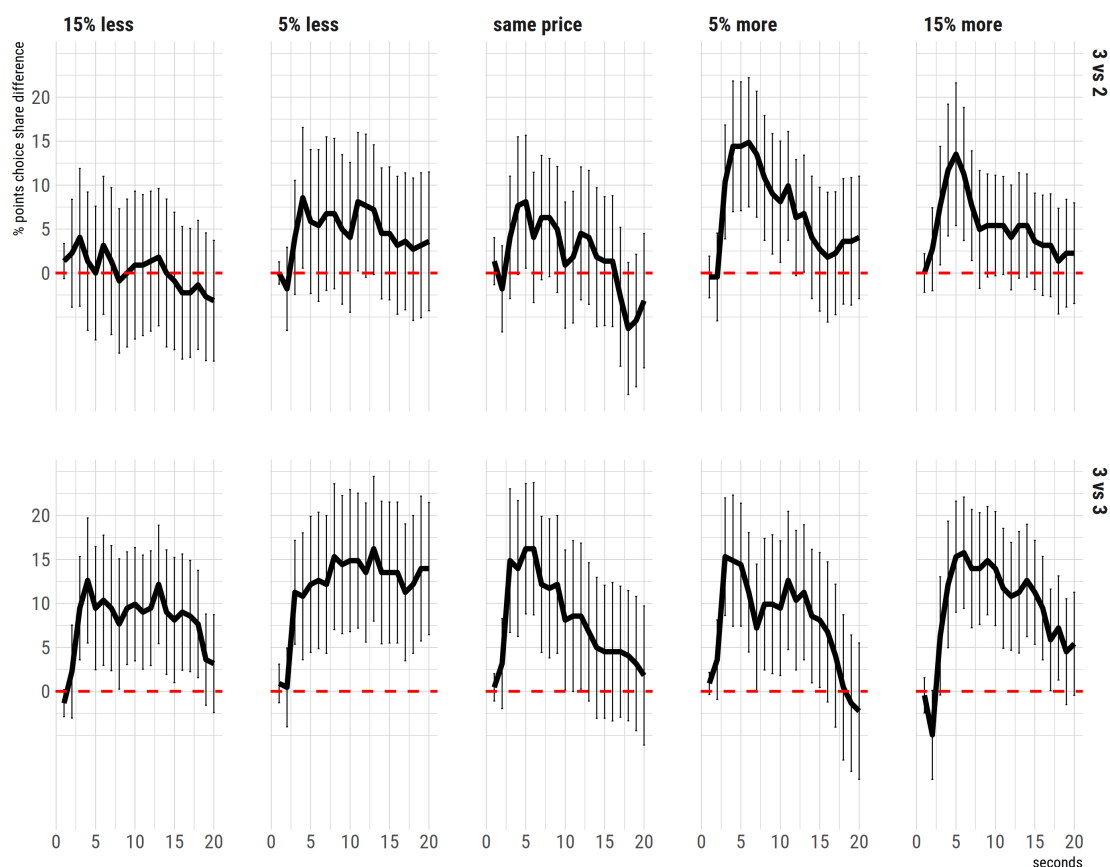


Figure 5: Measure of the AE, by markup and by type of menu comparisons.

Another way to look at the dynamic pattern of the AE is to look at the transition from the first to the second choices. Table 1 reports the choice shares of target, competitor and decoy in the first and second choice, across menus with three options without a decoy (left) and menus with three options including a decoy (right). In screens with a decoy, a strong preference for the target appears in the first choice; moreover, most people choosing the decoy move to the target when clicking for the second time. On the other hand, the breakdown of the second click is similar for target and competitor in screens without decoy, while it is biased in favor of the competitor in screens with a decoy. The pattern of clicks hence overall confirms that the target is disproportionately chosen in screens with a decoy in the early stages of the choice process, but then a revision process kicks in and participants switch away from the target, starting from the second click.

We now look at other aspects of the choice process, in particular, we ask how fast our participants made their choice, and how accurate their choices were on average. We present in figure 6 the share of participants not having made a choice, and of those having chosen the target, the competitor or the decoy, over time, across both treatments, for screens with three options, with and without a decoy. We see that more than 50% of choices are made before the 4th second, 50% of the rest are made before the 8th second, and so on. We also notice that choice appears to be faster in the graphical treatment, and that the choice share of the target appears to be lower in that treatment as well. However, it is not clear whether that is due to a lower AE or to lower precision, as indeed the choice share of the decoy is also higher in that treatment.

To track accuracy and the degree to which subjects respond to incentives, we present in figure 7 the

Screens without decoy				Screens with decoy			
First choice		Second choice		First choice		Second choice	
Target	43.4%	Competitor	26.9%	Target	55.3%	Competitor	36.0%
		Decoy	3.7%			Decoy	1.1%
		Stop	62.5%			Stop	57.1%
Competitor	47.8%	Target	28.2%	Competitor	37.0%	Target	27.6%
		Decoy	2.2%			Decoy	2.9%
		Stop	63.7%			Stop	62.2%
Decoy	8.3%	Target	41.3%	Decoy	7.0%	Target	63.4%
		Competitor	29.3%			Competitor	23.1%
		Stop	25.5%			Stop	7.1%
Stop	0.5%			Stop	0.7%		

Table 1: First choice and second choices as a function of first choices, in menus with three options, depending on the presence of a decoy.

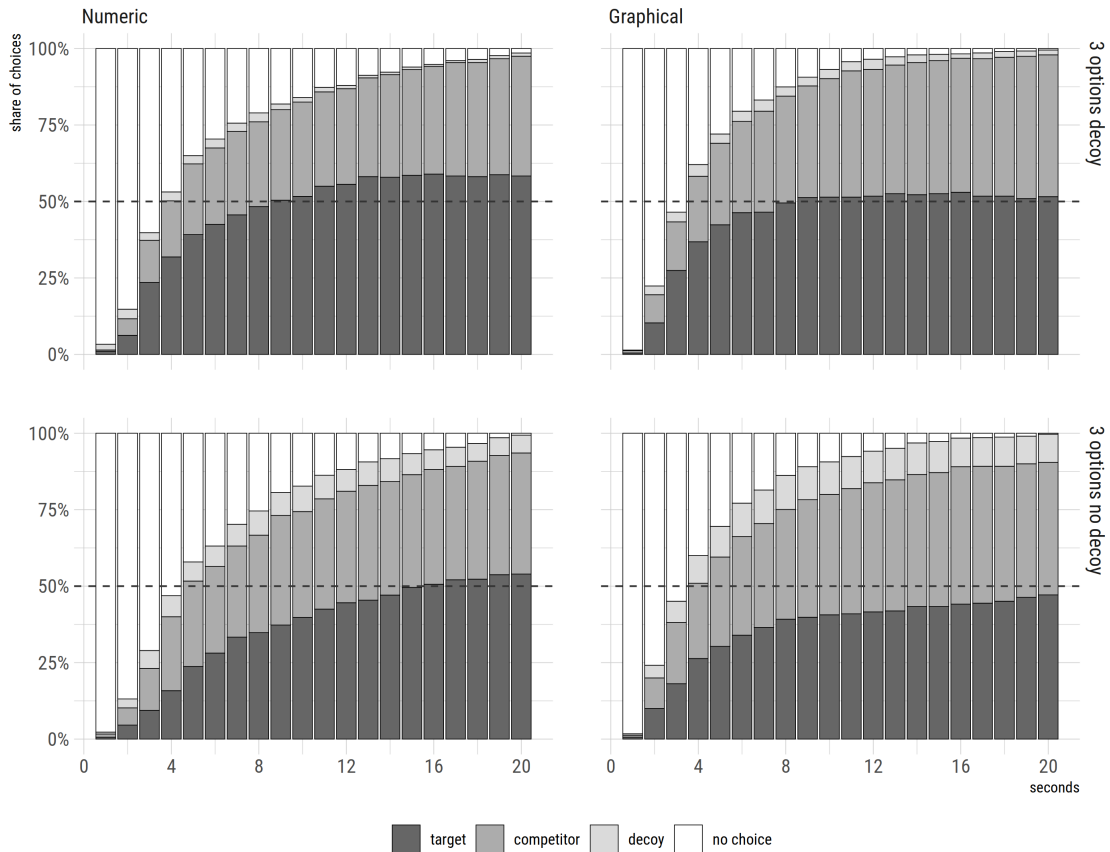


Figure 6: Choice shares, over time, by treatment, in menus with three options, depending on the presence of a decoy.

choice share of the target over time, across all screens, by treatment and by differences in the relative price of the target with respect to the competitor. The impact of the relative price of the target on its choice shares is particularly clear in the numeric treatment. This is probably because given time, the relative price could be computed. Differences are lower in the visual treatment, where the fuzzer stimuli may have induced more error. In both cases, though, the grey line representing the choice share of the target in case of indifference, i.e. when target and competitor have the same price, is significantly and consistently above 50%, replicating the usual findings of the literature, obtained in conditions of indifference. Finally, we see that even when the target was more expensive than the competitor, its choice share increases very fast at the beginning, to only stabilize or decrease later on. This points towards a choice process explanation, whereby participants exploit the dominance relation early on to avoid the decoy, and then go on to revise their choice by taking relative prices into account.

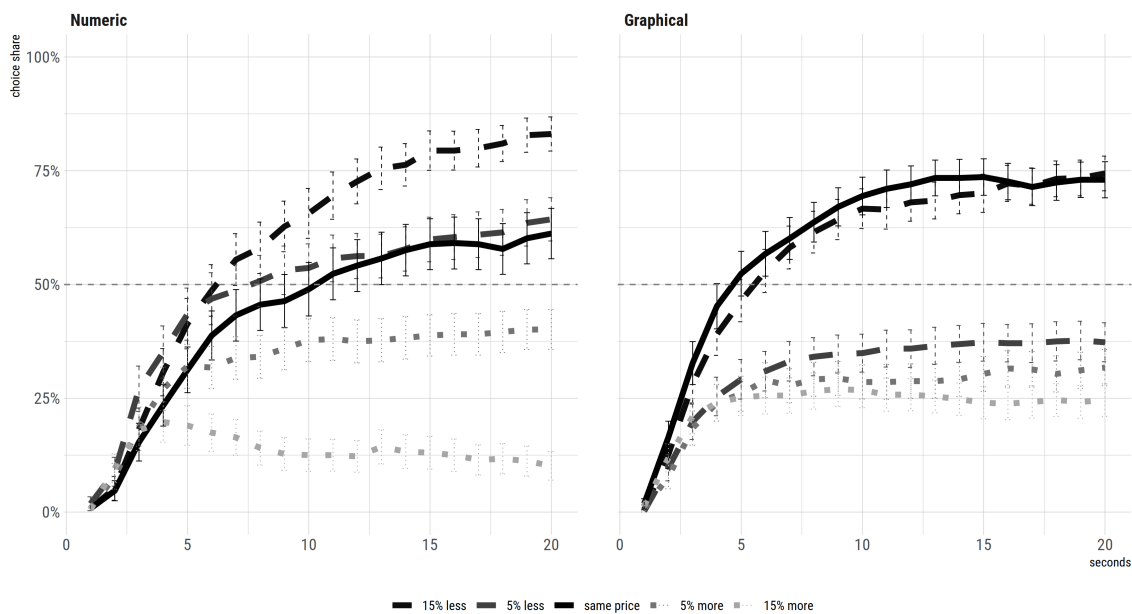


Figure 7: Choice share of the target by price difference.

## 5.2 A model of decision-making over time

In this section we model decision-making in the choice across options in menus with three options. This will allow us to separate different aspects of the dynamics of choices made in our experiment. In particular, we will focus not only on the strength of the AE over time, but also on the rise in the accuracy of choice, the extent to which participants avoid dominated options, and their willingness to make a choice rather than not selecting any option.

1. We assume that the utility of option  $x$  at time  $t$  is

$$u(x, t) = -\pi(x, t) \times p_x$$

with  $x$  the type of the option ( $x \in \{target, competitor, decoy\}$ ),  $t$  the time (between 0 and 20 seconds) and  $p_x$  the unit price of option  $x$ .

We will translate utilities into probabilities of choice by assuming that deciders perceive the utility of options with some error  $\epsilon_{xt}$ , i.i.d. across time and across options. We assume  $\epsilon_t$  is distributed



according to an extreme value distribution with precision parameter  $\lambda_t$  varying over time.

2. In menus with a decoy, we reflect the attraction effect and exclusion of dominated alternative by having two parameters translating the malus  $d$  applied to the decoy and the malus  $c$  applied to the competitor, over time  $t$ :

- $\pi(\text{decoy}, t) = d_t$
- $\pi(\text{target}, t) = 1$
- $\pi(\text{competitor}, t) = c_t$
- $\pi(\text{no\_choice}, t) = 1$

3. In menus without a decoy, then  $d_t = c_t = 1$ .

4. Finally, we consider the option of not making a choice, whereby:

- $u(\text{no\_choice}, t) = -v_t$

The no-choice option is therefore equivalent to a fourth option with implicit unit price  $v_t$ , to be estimated. We let  $v_t$  vary with time to translate the evolving willingness of consumers to make a choice or not. Since not making a choice results in a choice being made at random across options, a perfectly rational consumer who can evaluate all unit prices accurately would have  $v_t$  equal to the average of prices in the menu. That consumer would thus choose one option very quickly, as soon as he identified an obviously worse option.

Because a consumer cannot choose not to make a choice once he made a choice, we must recursively model the probability that no choice will have been expressed up to time  $t$ . Given the distribution of the error term  $\epsilon_{xt}$ , the probability not to make a choice in period  $t$ , conditional on not having made a choice in previous periods, is

$$p(\text{no\_choice}, t) = \frac{e^{-\lambda_t v_t}}{e^{-\lambda_t v_t} + \sum_x e^{\lambda_t u(x,t)}}$$

We then derive by recursion the probability  $P(\text{no\_choice}, t)$  not to have made a choice by period  $t$ :

$$P(\text{no\_choice}, t) = p(\text{no\_choice}, t) \times P(\text{no\_choice}, t - 1)$$

for  $t > 1$ , and we define  $P(\text{no\_choice}, 0) = 1$

Conditional on having made a choice, a consumer chooses option  $x \in \{\text{target}, \text{competitor}, \text{decoy}\}$  with probability:

$$p(x, t) = \frac{e^{\lambda_t u(x,t)}}{\sum_x e^{\lambda_t u(x,t)}} \quad (1)$$

At any point in time, therefore, option  $x$  other than *no\_choice* is thus chosen with probability  $(1 - P(\text{no\_choice}, t)) \times p(x, t)$ , and the no.choice option is chosen with probability  $P(\text{no\_choice}, t)$ .

### 5.3 Model estimation and results

We estimate parameters in our model for  $t \in \{4, 8, 12, 16, 20\}$ .<sup>2</sup> We first run Maximum-Likelihood estimation, then Bayesian estimation, for comparison with the maximum-likelihood results, and finally Bayesian estimation with random-effects (a.k.a. individual mixed effects). In those regressions, we

<sup>2</sup>Other divisions were also tested, but the main difference is between the first 10 seconds and the last 10 seconds, a difference which is captured by the division of time we present here.

allow our parameters to vary by individual and estimate the variability of those parameters in our sample. This allows us to more robustly estimate the mean of the parameters, abstracting from individual variations. Doing this allows us to use estimates of the parameters for broader inferences that are independent of the particular individuals in our sample. For each parameter  $param \in \{\lambda, c, d, v\}$ , we thus let  $param_{it} = param_t + \gamma_{it}$ , with  $i$  the individual and  $t$  the period. We estimate both  $param_t$  and the standard deviation of this parameter across individuals  $\gamma_{it}$ . Details on our Bayesian estimation model and procedure are given in appendix A. Table C.1 shows results of our estimation of parameters  $\lambda_t, bonus_t, malus_t$  and  $v_t$ , and of their standard deviation in the last set of regressions.

We find that Bayesian estimates without random effects are consistent with estimates from the Maximum-Likelihood estimation, which is what should happen as we use only weakly informative prior distributions for the parameters, meaning that we were agnostic about their value. Estimates of random effects show that there is a large variation in estimates across participants, which means that estimates without random effect are not reliable. We will therefore only discuss robust estimates of parameters from the model with mixed effects. We represent estimates of the Bayesian mixed effect estimations graphically in figure 8.

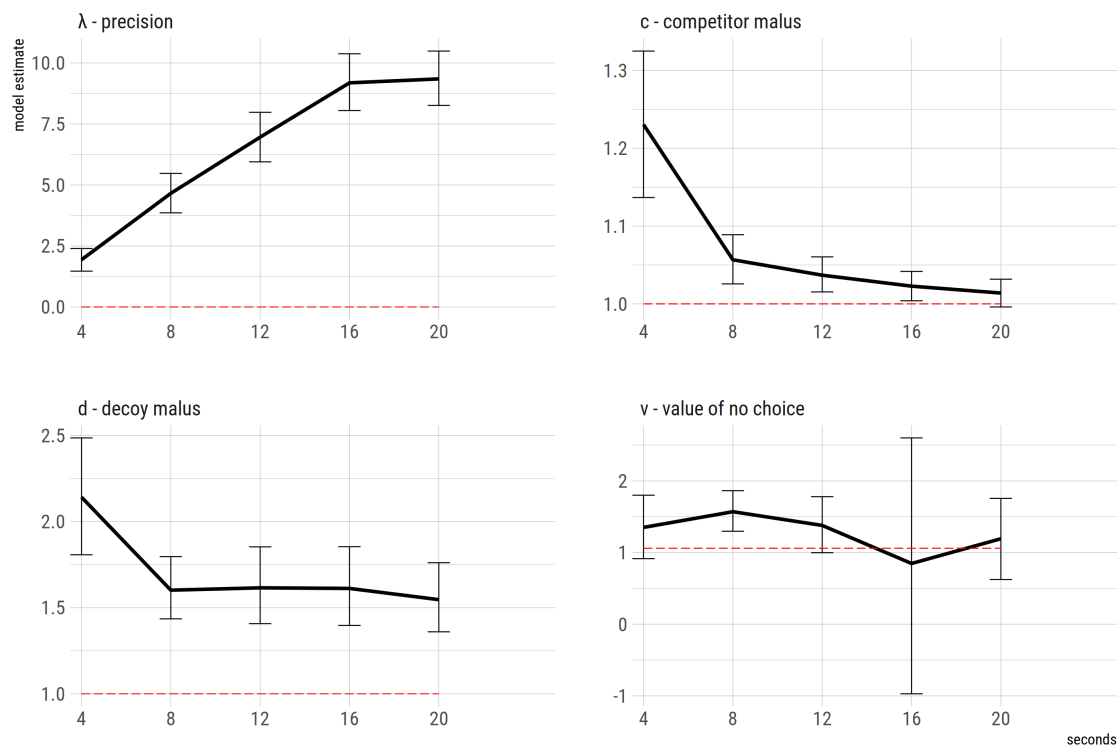


Figure 8: Parameter estimates, Bayesian model with mixed individual effects

Mean estimate as a line, 95% Bayesian credible intervals shaded.

The dashed reference lines are 0 for  $\lambda$ (choice at random), 1 for c and d (no bias) and 1.07 for  $v$  (average expected value of making a choice at random).

We observe that precision  $\lambda_t$  remains low after the first 4 seconds while parameters  $c_t$  and  $d_t$  are comparatively high. Later on, participants are more precise in their choice and both  $c_t$  and  $d_t$  become smaller. This means that participants are not able at first to decide precisely in terms of unit prices, but they compensate for this, when faced with menus with a decoy, by eliminating the decoy ( $d_t > 1$ ) and favoring the target vs. the competitor ( $c_t > 1$ ). This leads them to make early choices that favors the

target in excess of what would be rational given its unit price.

Estimates of  $c_t$  tell us that after 4 seconds, participants choose the target at a rate consistent with how often it would be chosen if there was no decoy and the competitor’s unit price was 24% higher ( $c_1 = 1.24$ ). This malus then declines to 6% after 8 seconds, 4% after 12 seconds, 2% after 16 seconds and 1% at the end. The 95% credible interval for  $c_t$  in periods 16 and 20 include the value 1, meaning that we cannot exclude the possibility that the target is not chosen significantly more often by the end of the time allocated for participants to make and revise their choices.

**Result 1.** *The attraction effect is more pronounced in the early stages of the decision process and then declines until the target is not significantly more often chosen than the competitor given their relative price.*

We finally consider the choice of not making a choice. Parameter  $v_t$  is relatively stable and higher than 1.07, which is the average expected price of an option chosen at random.<sup>3</sup> This means that participants behave as if they expect that a random choice — what they obtain as long as they do not make a choice — would obtain them an option with a price higher than what they can rationally expect. *Participants therefore prefer making a choice rather than making a choice at random.* The credible interval for  $v_t$  becomes wider with time. This is because of the low number of participants who still did not make a choice by the 16th or the 20th second. Later estimates are therefore based on only limited data, hence the lack of precision in their estimate.

### 5.3.1 Comparison across treatments

Table C.3 and Figure 9 present results of regressions outlining the difference in parameters depending on whether menus were presented graphically or numerically.

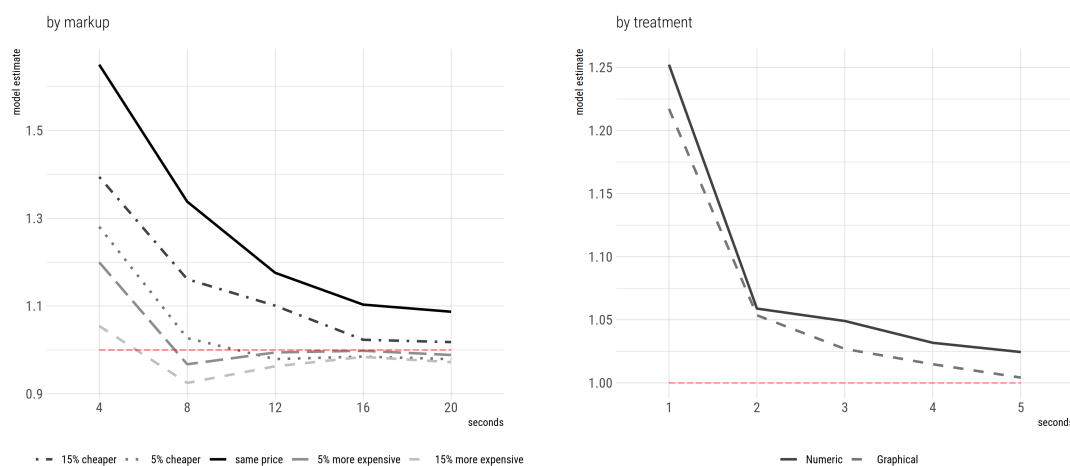


Figure 9: Mean parameter estimates, comparison by relative price of the target (left) and treatment (right), Bayesian model with mixed individual effects

We find that the attraction effect as measured by  $c_t$  is not significantly lower in the graphical treatment. The standard, numerical presentation does not therefore make the attraction effect higher, unlike stated in some recent literature. We however find that participants are more accurate in their choices in

<sup>3</sup>The decoy is priced on average at 1.2, the target at 1 and the competitor at 1.

the numerical treatment in the later part of the choice process. They are neither more likely to avoid the decoy or faster in making their choices. With reference to figure 6, our model allows us to understand that the lower share of the target in the graphical treatment is mainly due to lower precision in choice, which leads participants to both select the decoy more often, and to delay making a choice.

**Result 2.** *The attraction effect is not more pronounced when options are presented numerically rather than graphically.*

### 5.3.2 Comparison by markup

Table C.4 and Figure 9 presents results controlling for the relative price of the target with respect to the competitor. The markup varied across menus from 15%, whereby the target is more expensive than the competitor by 15%, to -15%. We estimate  $\lambda$ ,  $c$ ,  $d$  and  $v$  across all menus. Our baseline results are shown in column 1, where we show the case where the target and the competitor are of the same price (markup=0%). We find there that while the attraction effect  $c$  decreases over time, it is still significantly more than 1 by the time limit for making decisions ( $c_5$  is on average 1.08 with a 95% credible interval of [1.05,1.11]). Using our perspective, this is because there is no gain in exerting effort beyond eliminating the decoy if the target and the competitor are equivalent. In that case, there is no loss in favoring the target. Columns 2 to 5 show by how much  $c$  differs from the baseline when the markup is different from 0. We find that the attraction effect is significantly less pronounced in those cases, as was hypothesized.

**Result 3.** *The attraction effect is more pronounced when the choice across options is indifferent.*

This confirms the statement in Huber et al. (2014) whereby the attraction effect is more likely to emerge in cases where there is no clear difference between options. In those case, it does not matter that much of the literature on the attraction effect relies on experiments that are not incentivized, since there is no gain in going beyond eliminating the decoy. However, this result shows how important it is to incentivize choice in order to investigate the practical and theoretical significance of the attraction effect.

## 5.4 Individual differences in decision modes

The analysis shown above relies on the implicit assumption that all participants follow the same strategy. At the aggregate level, our participants behave *as if* they switched from a heuristic-driven decision style based on dominance in the first seconds of the choice process, to an accurate assessment of the options and a maximizing decision style in the remainder of their allotted time. This aggregate behavior gives rise to the rise-and-fall pattern that we described and modeled in the previous sections.

Nonetheless, the observed pattern could be due to the simultaneous existence in the population of a mix of different types of decision-makers. For example, it might be that some participants follow heuristics and choose fast, while other compare all options and choose based on their perception of their unit price, which makes their choice slower. We would then also observe fast, heuristic driven choice favoring the target at the beginning, and slow, more precise and unbiased choice, at the end. This would be a composition effect of different participants, whereby subject would not be switching decision modes, as hypothesized by us, but rather differ in decision models.

We find that our participants do indeed display a high variety in their response patterns. Figure 10 shows the choice patterns of nine participants, using the same basic plot introduced for the aggregate results – the difference in the choice share of the target between screens with and without a decoy at each point in time. For simplicity we report only the  $3vs3$  measure.

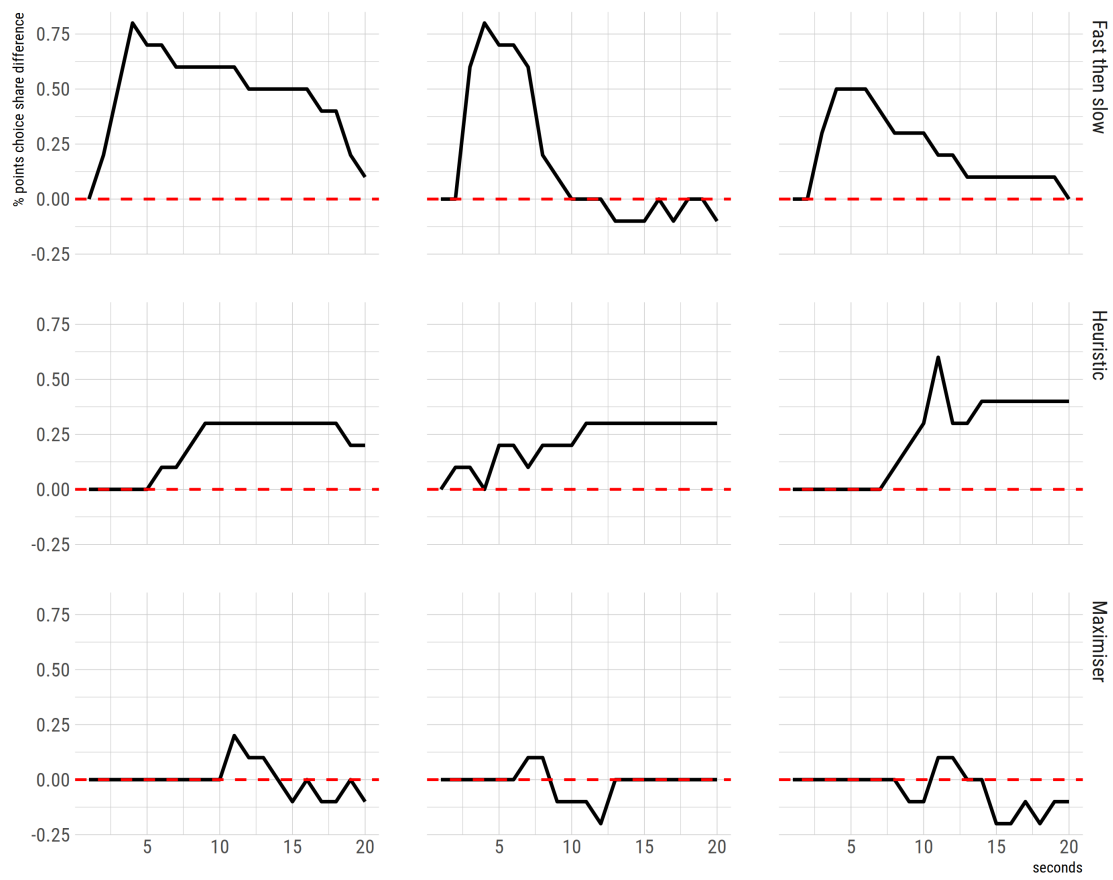


Figure 10: Types of participants, 3vs3 measure of the AE

The first three participants are examples of what we call the *fast then slow* type. They rely extensively on dominance in the first seconds – the choice share of the target increases significantly in presence of a decoy – but then they revise their choices to get at the end of the allotted time at or near the point of no extra preference for the target. These participants replicate at the individual level the same pattern as the one we found at the aggregate level. The participants in the second row are examples of *heuristic-following* participants. They might take some time to submit a decision (horizontal line) or not, but in all cases prefer the target also in later stages of the choice process, and continue to do so as the time runs out. They sooner or later give a preference to the target in presence of a decoy, and they never revise their choice. The participants in the third row are what we call *maximisers*. They wait to submit their first choice, and when they do, they do not particularly show a preference for the target. They follow their assessment of unit prices only, disregarding dominance, and end up near or on the point of no extra preference for the target.

While not all participants show patterns that are consistent with these three types, most do. Subjects respond differently to dominance, and while the AE does appear at the aggregate level as an intuitive first response to then decline, its relative importance is indeed different across participants.

#### 5.4.1 Estimates from a mixture model

In a last section of our analysis, we therefore explore the possibility that our results may be driven by the existence of a mix of different type of decision-makers in the population. We interpreted our results up to now as driven by a process by which consumers switch from a heuristic-driven style of decision making to a maximizing style after the decoy option has been eliminated. However, the same pattern of rise and fall of the attraction effect could be the result of different patterns of choice across participants, some following heuristics and choosing fast, the others comparing all options and choosing based on their perception of their unit price, which makes their choice slower. We would then also observe fast, heuristic driven choice at the beginning, and slow, more precise choice, at the end.

To explore this possible explanation for our result, we define three types of decision-makers<sup>4</sup> as follows:

- HM decision makers first follow a **heuristic** and then **maximize**. They decide as outlined in section 5.2. Parameters to estimate for this type of decision maker are  $\lambda_t, c_t, d_t$  and  $v_t$ .
- M decision makers are **maximizers**. For those, we restrict  $c_t = d_t = 0$ . This means that they take account only of the unit price and are neither subject to the attraction effect or take account of dominance. Parameters to estimate are then  $\lambda_t$  and  $v_t$ .
- H decision makers follow a **heuristic**. Those choose based only on characteristics of the offers without re-evaluating their decision over time. That is, all parameters are constant over time except for  $v_t$ . Those decision-makers take unit prices into account, avoid the decoy and favor the target but do not take advantage of time to increase the precision of their choice.

We estimate the proportion of each type in our population by running a mixture model on our data (Figure 11 and Table C.5 in appendix B).

Results from the Bayesian mixture model with mixed effects indicate that about 61% of our participants are HM decision-makers, that is, they switch from a fast to a slow decision process. Those participants increase precision  $\lambda$  over time while the attraction effect  $c$  decreases over time until not being significantly different from 1. About 18% of the participants are maximizers, that is, they are not

<sup>4</sup>We also considered the possibility of some decision-makers choosing at random ( $\lambda = 0$ ), but estimated the proportion of such random choosers to be less than 1%.

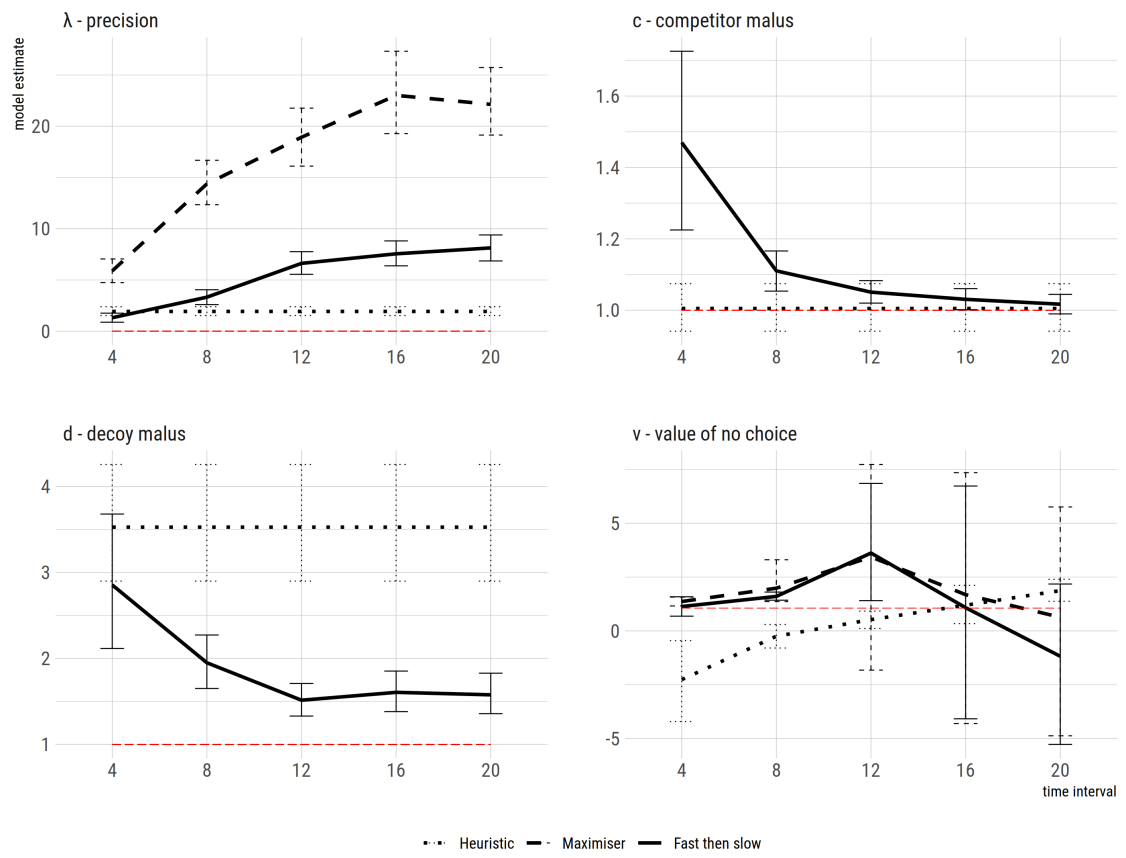


Figure 11: Comparison of parameters, by type.

subject to the attraction effect and do not take advantage of the dominance relation beyond what is predicted from the difference in price between the decoy and the target. Maximizers are more precise than HM or H decision-makers, and their precision increases over time. Finally, about 21% of the participants are heuristic decision-makers. Those participants who follow heuristics are the most reluctant to make a choice (low values of parameter  $v$ ), are also less precise than others (low  $\lambda$ ), avoid the decoy ( $d > 1$ ) but are not attracted by the target ( $c$  not significantly different from 1). It might be that those participants are those who are most confused by the choice to be made, and therefore both delay choice and rely on a simple decoy elimination heuristics.

**Explaining types.** The classification of types that is obtained from the Bayesian mixture model is well defined, in the sense that the vast majority of individuals are unequivocally classified as either HM, H or M decision makers. In particular, the distinction between types H and M is very definite, as there are no individuals that may be either H or M (Figure D.1 in appendix C).

We perform multinomial logistic regressions of the likelihood that an individual belongs to a type, as a function of the treatment (numeric or graphical) and of individual’s characteristics. Those were elicited from our preliminary questionnaire including shopping habits, proneness to confusion, risk and loss aversion, as well as the individual’s CRT scores and socio-demographic characteristics (see section 4, further description of individual statistics in appendix D and appendix E for summary statistics). In our regressions, type HM is the baseline category. We select variables by the Akaike Information Criterion in a stepwise algorithm, adding and removing terms until a maximum is achieved. Results are shown in table 2 where we see if types H and M differ significantly from type HM along different variable. In the first two columns, the type of an individual is the one he is most likely to belong to according to our Bayesian mixture model (mode). In the last two columns, we allow uncertainty in the assignment of types, so the dependent variable is  $(p_{HM,i}, p_{H,i}, p_{M,i})$ . Results are consistent across both specifications.

	Modal Type		Mean type	
	H	M	H	M
CRT score	-1.0** [-1.7, -0.2]	-0.1 [-0.6, 0.5]	-0.8** [-1.6, -0.1]	-0.1 [-0.6, 0.5]
Risk tolerance	0.3** [0.005, 0.6]	0.2 [-0.1, 0.4]	0.2 [-0.05, 0.5]	0.1 [-0.2, 0.4]
1.Student	0.2 [-2.2, 2.7]	-2.7** [-5.0, -0.4]	0.4 [-1.9, 2.8]	-2.2* [-4.4, 0.04]
1.Worker	1.8 [-0.6, 4.2]	-1.2 [-3.6, 1.2]	1.7 [-0.7, 4.0]	-1.0 [-3.4, 1.5]
Education level	0.2 [-0.3, 0.7]	1.1**[0.3, 1.9]	0.1 [-0.4, 0.6]	1.1***[0.3, 1.9]
Econ student	1.8** [0.4, 3.1]	-0.6 [-2.0, 0.8]	1.4** [0.1, 2.7]	-0.8 [-2.2, 0.6]
Age	0.1** [0.003, 0.2]	-0.2 [-0.4, 0.1]	0.1** [0.005, 0.2]	-0.2* [-0.5, 0.03]
Revenue	-0.9 [-2.7, 1.0]	-2.5** [-4.7, -0.3]	-0.9 [-2.8, 0.9]	-2.3** [-4.5, -0.04]
Graphical treatment	-0.2 [-1.4, 1.0]	-1.9***[-3.2, -0.5]	-0.5 [-1.6, 0.7]	-2.0***[-3.3, -0.6]
Constant	-6.4***[-11.2, -1.6]	1.9 [-4.4, 8.1]	-5.6** [-10.1, -1.1]	2.5 [-3.9, 8.8]
Akaike Inf. Crit.	190.0	190.0	192.9	192.9

Note: 95% CI in parenthesis.\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

Table 2: Results of multinomial regressions of type on individual characteristics

We find that there are more maximizers in the numeric treatment (29% vs. 10%,  $p < 1\%$ ). This might be because a numeric presentation of options allows participants to rely on unit-price comparisons by computing price over quantity ratios. We also find that individuals who score lower in the CRT are more likely to be of type H. This is consistent with the view that wrong answers to the CRT correspond to an intuitive decision mode. People who are inactive (that is, neither students or workers), and those who are less well-off in terms of revenues, are more likely to be of type M. This points towards the motivation to do well monetarily in the experiment as a drivers towards maximizing behavior. Those who have higher education levels are also more likely to be of type M, meaning that cognitive ability may also play a role in adopting that strategy.



## 6 Discussion

We provide choice-process, individual-level evidence that the attraction effect is not a persistent feature of choice but rather a temporary, shortlisting heuristic that drives choice in the first seconds of exposure to stimuli but largely disappears in the long run. We show this to be true at the aggregate level but also for a majority share of participants at the individual level. Most participants follow what we call a *fast then slow* strategy: they employ a fast and frugal heuristics in the first five to ten seconds of exposure to the task, and then revise their choice leading to changes in the direction of optimal behavior.

Considering the attraction effect a simple, fast decision strategy allows to rationalize several disconnected results from the literature and shed light on the recent controversies about the applicability of the AE in real consumer decisions. The large list of settings in which the AE fails can be grouped in two different strands: settings in which other heuristics allow the participants to do better; and settings where the subject is incentivized to switch away from heuristic decision making and towards “slow”, “reflective” and “maximizing” strategies. Switch to competing heuristics more appropriate to the problem at hand can easily explain why the AE is muted in presence of other focal cues (like brand), or when the product description is detailed and unambiguous, or when the products are known and familiar to the consumer. Difficulty in recognizing and being relatively certain of dominance can explain the muting of the AE with pictorial representation of products and in real-world choices. Switching to “System 2”, slow decision modes can explain the disappearance of the effect away from indifference.

The view of the AE as a simple dominance-based heuristic is backed by several other existing results in the literature. [Mao and Oppewal \(2012\)](#) showed the AE to be more pronounced among consumers with an intuitive thinking style as measured by an abridged version of the Rational-Experiential Inventory questionnaire ([Epstein et al., 1996](#)). [Pocheptsova et al. \(2009\)](#) showed that the AE increases when participants’ cognitive resources were depleted and ([Masicampo and Baumeister, 2008](#)) when participants were tired. [Hu and Yu \(2014\)](#) showed in an fMRI study that the AE was associated with activation in areas of the brain linked with heuristics, and that those participants who had lower AE had activation in areas linked to cognitive control. [Howes et al. \(2016\)](#) show that context effects, as heuristics in general, can be optimal when signals are noisy or decision makers are inaccurate in their assessment of options. On the other hand, our results contradict the findings of both [Pettibone \(2012\)](#) on the AE and [Dhar et al. \(2000\)](#) on the compromise effect, who show that increasing time pressure decreases the size of context effects. We think that our design is more robust in several ways – those two papers used non-incentivized choices, ran between subjects analyses, and imposed indifference between options – thus yielding stronger results.

More generally, our findings show that participants can switch between decision modes when given incentives to do so. This point was made already in early literature on the topic ([Payne et al., 1988, 1993](#)), that argued that participants use phased decision strategies, whereby they employ different types of processing at different phases of the decision. Our experiment allows us to provide some clear, incentivized and measurable empirical evidence behind the use of such phased decision processes. Our findings also resonate with a very recent literature on the possibility to correct behavioral biases by confronting participants with their inconsistencies and offering them to revise their submitted choices ([Nielsen and Rehbeck, 2019](#)).

It is possible that our design creates artefactual incentives for participants to adopt a *fast then slow* heuristic. Indeed, our choice process elicitation mechanism incentivizes rational decision makers that are aware of their cognitive limits to provide a first, intuitive reply and then revise it. If a rational decision maker could make a fast, accurate first reply she should do it. But if she knows that she needs time to provide an accurate choice, then it is in her best interest to submit a provisional choice that would allow her to avoid the dominated option and then revise her choice. We think that this

artefactual effect should be a minor concern. First, this artifact should not decrease the external validity of our findings. In real life participants could use the exact same reasoning – provide an intuitive reply in case of limited time and cognitive resources, think through the choice if time and incentives allow for it; our design just allows us to observe both strategies at the same time, and the switching point between the two. Second, despite the incentive to be *fast then slow*, a sizable share of participants preferred to wait and maximise – i.e. not to exploit dominance - and a similar share of participants chose to exploit only dominance and forfeit further adjustments. This shows us that participants can indeed follow their preferred choice path within our incentive scheme.

Our design focuses solely on the attraction effect, but it could be in principle applied to any context effect, by choosing appropriate stimuli. The method scales easily, provides individual estimates of behavioral biases and exposes the choice process at play in participants' minds. If applied to other context effects or behavioral biases, it could provide a contribution to the debates around the persistence of behavioral effect, their theoretical nature, and the limits of rational choice.

If the results would replicate across different context effects, the importance of the context effect literature as a powerful violation of rational choice would be diminished, as the core axioms of rationality would be shown to hold when participants are given time and incentives to ponder their responses. The relevant empirical question would then become to understand under which conditions participants choose different decision modes, and hence in which contexts should we expect the behavioral biases identified in the growing behavioral literature to apply.

Our findings about the short-term nature of the attraction effect help to make sense and generalize the debate in marketing on the existence and robustness of the AE, and provide a guide to assess the potential real-world implications of this bias.

In theoretical terms, thinking of the AE as a short-term decision strategy limits its role in reshaping existing theories of rational choice; as most heuristics, it is locally rational, and if applied in its proper context and in presence of bounded rationality, limited accuracy or noisy information it improves pay-offs; as most heuristics, it does not invalidate more deductive and comprehensive accounts of rational choice theories outside of its limited scope. The IIA axiom is alive and well given enough time and incentives. After a few seconds, all but a minority share of heuristic-only decision makers switch to more reflective and analytic decision modes, thus abandoning the provisional AE-driven choice and exploiting their accuracy to make better informed choices. This is true for all situations away from indifference. On the other hand, in conditions of indifference among options, the effect persists until the end of the allotted time. This is because in conditions of indifference, any cue might be used to make a choice, and dominance provides an easy tie-breaking rule.

In terms of the practical applicability of AE research, our results imply that the AE is bound to have a small to negligible effect on choice – bar on the few situations in which dominance really stands out. First, the AE faces competition from other heuristics. In most applied contexts, it is likely to be washed out by other cues, as familiarity, brand preferences, other, non-dominance based salient cues, and habits. Second, in most contexts it is hard to put the AE to use. For the AE to work, the dominance relation must be unambiguous and focal. In most experiments the dominance relationship is artificially enhanced to be central, clear, and unmissable. This is likely not to be the case in applied settings.

We thereby counter the present focus on distinguishing between the types of context or the type of person following this or that decision mode by showing that many people are able to switch from one to the other decision mode if given the time to do so. Our work is thus a contribution to the interpretation of context effects and to the epistemic debate surrounding those, and also a contribution to the literature on the dynamics of choice processes. We help to make sense and generalize the debate in marketing on the existence of the AE. Finally, our work also has real-world implications, in that context effects may not be relevant when consumers have sufficient time and motivation to choose the best option in

a menu. This limits the extent to which the AE can be exploited by marketers in order to influence consumer choice.

## References

- Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*, 129(1):139–167.
- Ariely, D. and Wallsten, T. S. (1995). Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect. *Organizational Behavior and Human Decision Processes*, 63(3):223–232.
- Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.
- Brooks, S. P. and Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.
- Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7):2899–2922.
- Castillo. Testing the asymmetric dominance effect and its explanations.
- Chang, S.-S., Chang, C.-C., and Liao, Y.-Y. (2015). A joint examination of effects of decision task type and construal level on the attraction effect. *Journal of Economic Psychology*, 51:168–182.
- Cohen, P. M. and Santos, L. R. (2017). Capuchins (*cebus apella*) fail to show an asymmetric dominance effect. *Animal Cognition*, 20(2):331–345.
- Crosetto, P. and Gaudeul, A. (2016). A monetary measure of the strength and robustness of the attraction effect. *Economics Letters*, 149:38–43.
- Denwood, M. J. (2016). runjags: An R Package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software*, 71(9).
- Dhar, R. (1996). The Effect of Decision Strategy on Deciding to Defer Choice. *Journal of Behavioral Decision Making*, 9(4):265–281.
- Dhar, R. and Nowlis, S. M. (1999). The Effect of Time Pressure on Consumer Choice Deferral. *Journal of Consumer Research*, 25(4):369–384.
- Dhar, R., Nowlis, S. M., and Sherman, S. J. (2000). Trying Hard or Hardly Trying: An Analysis of Context Effects in Choice. *Journal of Consumer Psychology*, 9(4):189–200.
- Dieckmann, A., Dippold, K., and Dietrich, H. (2009). Compensatory versus noncompensatory models for predicting consumer preferences. *Judgment and Decision Making*, 4(3):200–213.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2008). Representative Trust and Reciprocity: Prevalence and Determinants. *Economic Inquiry*, 46(1):84–90.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of personality and social psychology*, 71(2):390.
- Farmer, G. D., Warren, P. A., El-Deredy, W., and Howes, A. (2016). The Effect of Expected Value on Attraction Effect Preference Reversals. *Journal of Behavioral Decision Making*, page n/a–n/a.

- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4):25–42.
- Frederick, S., Lee, L., and Baskin, E. (2014). The Limits of Attraction. *Journal of Marketing Research*, 51(4):487–507.
- Gigerenzer, G., Todd, P. M., and Gerd Gigerenzer Group, A. R. (2000). *Simple Heuristics That Make Us Smart*. Oxford University Press, Oxford, 1 edition edition.
- Hedgcock, W. and Rao, A. R. (2009). Trade-off aversion as an explanation for the attraction effect: A functional magnetic resonance imaging study. *Journal of Marketing Research*, 46(1):1–13.
- Herne, K. (1999). The effects of decoy gambles on individual choice. *Experimental Economics*, 2(1):31–40.
- Howes, A., Warren, P. A., Farmer, G., El-Deredy, W., and Lewis, R. L. (2016). Why contextual preference reversals maximize expected value. *Psychological Review*, 123(4):368–391.
- Hu, J. and Yu, R. (2014). The neural correlates of the decoy effect in decisions. *Frontiers in Behavioral Neuroscience*, 8:271.
- Huber, J., Payne, J. W., and Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research*, 9(1):90–98. 01119.
- Huber, J., Payne, J. W., and Puto, C. P. (2014). Let's be honest about the attraction effect. *Journal of Marketing Research*, 51(4):520–525.
- Huber, J. and Puto, C. (1983). Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects. *Journal of Consumer Research*, 10(1):31–41.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Król, M. and Król, M. (2019). Inferiority, not similarity of the decoy to target, is what drives the transfer of attention underlying the attraction effect: Evidence from an eye-tracking study with real choices. *Journal of Neuroscience, Psychology, and Economics*, 12(2):88.
- Kroll, E. B. and Vogt, B. (2012). The relevance of irrelevant alternatives. *Economics Letters*, 115(3):435–437.
- Lichters, M., Bengart, P., Sarstedt, M., and Vogt, B. (2015). What really matters in attraction effect research: when choices have economic consequences. *Marketing Letters*, page 1–12.
- Lohse, G. L. and Johnson, E. J. (1996). A Comparison of Two Process Tracing Methods for Choice Tasks. *Organizational Behavior and Human Decision Processes*, 68(1):28–43.
- Malkoc, S. A., Hedgcock, W., and Hoeffler, S. (2013). Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives. *Journal of Consumer Psychology*, 23(3):317–329.
- Mao, W. and Oppewal, H. (2012). The attraction effect is more pronounced for consumers who rely on intuitive reasoning. *Marketing Letters*, 23(1):339–351.
- Masicampo, E. and Baumeister, R. F. (2008). Toward a Physiology of Dual-Process Reasoning and Judgment: Lemonade, Willpower, and Expensive Rule-Based Analysis. *Psychological Science*, 19(3):255–260.
- Mishra, S., Umesh, U., and Stem Jr, D. E. (1993). Antecedents of the attraction effect: An information-processing approach. *Journal of Marketing Research*, 30(3):331–349.
- Nielsen, K. and Rehbeck, J. (2019). Are axioms normative? eliciting axiom preferences and resolving conflicts with lottery choices. Working paper.
- Park, J. and Kim, J. (2005). The effects of decoys on preference shifts: The role of attractiveness and providing justification. *Journal of Consumer Psychology*, 15(2):94–107.

- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3):534–552.
- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.
- Pettibone, J. C. (2012). Testing the effect of time pressure on asymmetric dominance and compromise decoys in choice. *Judgment and Decision Making*, 7(4):513.
- Pocheptsova, A., Amir, O., Dhar, R., and Baumeister, R. F. (2009). Deciding without resources: Resource depletion and choice in context. *Journal of Marketing Research*, 46(3):344–355.
- Ratneshwar, S., Shocker, A. D., and Stewart, D. W. (1987). Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity. *Journal of Consumer Research*, 13(4):520–533.
- Ray, D. and Robson, A. (2018). Certified Random: A New Order for Coauthorship. *American Economic Review*, 108(2):489–520.
- Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study. *American Economic Review*, 101(2):900–926.
- Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological review*, 108(2):370.
- Schuck-Paim, C., Pompilio, L., and Kacelnik, A. (2004). State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS Biol*, 2(12):2305–2315.
- Shafir, S., Waite, T. A., and Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*apis mellifera*) and gray jays (*perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, 51(2):180–187.
- Simon, H. A. (1959). Theories of Decision-Making in Economics and Behavioral Science. *The American Economic Review*, 49(3):253–283.
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16(2):158–174.
- Sonsino, D. (2010). The irrelevant-menu affect on valuation. *Experimental Economics*, 13(3):309–333.
- Spiliopoulos, L. and Ortmann, A. (2017). The BCD of response time analysis in experimental economics. *Experimental Economics*, page 1–51.
- Sürücü, O., Djawadi, B. M., and Recker, S. (2019). The asymmetric dominance effect: Reexamination and extension in risky choice – an experimental study. *Journal of Economic Psychology*, 73:102–122.
- Trendl, A., Stewart, N., and Mullett, T. (2018). A zero attraction effect in naturalistic choice. *SSRN Electronic Journal*.
- Trueblood, J. S., Brown, S. D., Heathcote, A., and Busemeyer, J. R. (2013). Not just for consumers context effects are fundamental to decision making. *Psychological science*, 24(6):901–908.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79(4):281–299.
- Tversky, A. (2003). Elimination by Aspects: A Theory of Choice. *Preference, Belief, and Similarity: selected writings*, 463.
- Tversky, A. and Simonson, I. (1993). Context-Dependent Preferences. *Management Science*, 39(10):1179–1189. 00885.
- Usher, M. and McClelland (2004). Loss aversion and inhibition in dynamical models of multi-alternative choice. *Psychological Review*, 111(3):757–769.
- Walsh, G., Hennig-Thurau, T., and Mitchell, V.-W. (2007). Consumer confusion proneness: scale development, validation, and application. *Journal of Marketing Management*, 23(7-8):697–721.

- Wang, Z., Jusup, M., Shi, L., Lee, J.-H., Iwasa, Y., and Boccaletti, S. (2018). Exploiting a cognitive bias promotes cooperation in social dilemma experiments. *Nature communications*, 9(1):2954.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1):67-85.
- Yang, S. and Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research*, 51(4):508-513.

## A Bayesian model

### Likelihood

Denote *tar* the target, *com* the competitor, *dec* the decoy and *nc* the no choice option.  $p_{tar}$ ,  $p_{com}$  and  $p_{dec}$  are the price of three options in a menu,  $1_{dec}$  is a binary variable taking value 1 if the menu includes a decoy,  $\lambda_t, c_t, d_t, v_t$  are the parameters of our model. Then we compute perceived utility  $per_{x,t}$  of each option  $x$  at time  $t$  as follows:

$$\begin{aligned} per_{tar,t} &= \exp(-\lambda_t \times p_{tar}) \\ per_{com,t} &= \exp(-\lambda_t \times p_{tar} \times (1 - 1_{dec} + c_t \times 1_{dec})) \\ per_{dec,t} &= \exp(-\lambda_t \times p_{dec} \times (1 - 1_{dec} + d_t \times 1_{dec})) \\ per_{nc,t} &= \exp(-\lambda_t \times v_t) \end{aligned}$$

Denote  $prob_{x,t}$  for options  $x \in \{target, comp, decoy\}$  as

$$prob_{x,t} = per_{x,t} / \sum_{x \in \{target, comp, decoy\}} per_{x,t}$$

and denote  $prob_{nc,t}$  as

$$prob_{nc,t} = per_{nc,t} / \sum_{x \in \{target, comp, decoy, nc\}} per_{x,t}$$

Then the probability  $pr_{nc,t}$  that option *nc* is chosen at time  $t$  is such that

$$\begin{aligned} pr_{nc,1} &= prob_{nc,1} \\ pr_{nc,t} &= prob_{nc,t} \times pr_{nc,t-1} \end{aligned}$$

and the probability  $pr_{x,t}$  that option  $x \in \{tar, com, dec\}$  is chosen at time  $t$  is

$$pr_{x,t} = (1 - pr_{nc,t}) \times prob_{x,t}$$

Finally, the likelihood option  $y_t$  is chosen at time  $t$  follows a categorical distribution over the four options:

$$y_t \sim \text{Cat}(pr_{tar,t}, pr_{com,t}, pr_{dec,t}, pr_{nc,t})$$

### Mixture model

In the mixture model, we assume there are three types of people, type HM with parameters to be estimated as in section A, type M with parameters  $c_t, d_t$  restricted to be equal to 0, and H with parameters  $\lambda_t, c_t, d_t$  restricted to be constant over time. The likelihood individual  $i$  belong to type HM, H or M follows a categorical distribution over the three options:

$$type_i \sim \text{Cat}(p)$$

with  $p = (pr_{HM}, pr_H, pr_M)$

## Priors

We use the following weakly informative prior distributions:  $\lambda_t \sim \Gamma(0.1, 0.1)$ ,  $c_t \sim \mathcal{N}(1, 0.1)$ ,  $d_t \sim \mathcal{N}(1, 0.1)$ ,  $v_t \sim \mathcal{N}(1, 0.1)$ .

$\Gamma(\alpha, \beta)$  denotes the gamma distribution with share  $\alpha$  and rate  $\beta$ . Its mean is  $\alpha/\beta$  and its standard deviation is  $\sqrt{\alpha}/\beta$ .

$\mathcal{N}(\alpha, \beta)$  is the normal distribution with mean  $\alpha$  and precision  $\beta$ . Its standard deviation is  $\sigma = \sqrt{1/\beta}$ .

Both  $\Gamma(0.1, 0.1)$  and  $\mathcal{N}(1, 0.1)$  therefore have mean 1 and standard deviation  $\sqrt{10}$ .

## Priors for hierarchical model

In our extension of the model to allow for mixed effects, we use the following prior distributions:  $\lambda_{it} \sim \mathcal{N}(\lambda_t, \tau_{\lambda,t})$ ,  $c_{i,t} \sim \mathcal{N}(c_t, \tau_{c,t})$ ,  $d_{i,t} \sim \mathcal{N}(d_t, \tau_{d,t})$ ,  $v_{it} \sim \mathcal{N}(v_t, \tau_{v,t})$ , while  $\lambda_t \sim \Gamma(0.1, 0.1)$ ,  $c_t \sim \mathcal{N}(1, 0.1)$ ,  $d_t \sim \mathcal{N}(1, 0.1)$ ,  $v_t \sim \mathcal{N}(1, 0.1)$  as before and  $\tau_{\lambda,t} \sim \Gamma(0.1, 0.1)$ ,  $\tau_{c,t} \sim \Gamma(0.1, 0.1)$ ,  $\tau_{d,t} \sim \Gamma(0.1, 0.1)$ ,  $\tau_{v,t} \sim \Gamma(0.1, 0.1)$ .

## Priors for mixture model

The prior for  $p$  is a Dirichlet distribution  $Dir(1, 1, 1)$ . We use weakly informative priors for  $\lambda_{type,t}$ ,  $c_{type,t}$ ,  $d_{type,t}$ ,  $v_{type,t}$  of each type HM, M or H, whereby  $\lambda_{type,t} \sim \Gamma(0.1, 0.1)$ ,  $c_{type,t} \sim \mathcal{N}(1, 0.1)$ ,  $d_{type,t} \sim \mathcal{N}(1, 0.1)$ ,  $v_{type,t} \sim \mathcal{N}(1, 0.1)$ .

## Estimation

We run JAGS 4.3.0 on R 3.5.3 with runjags 2.0.4 to estimate Bayesian models (Denwood, 2016). Estimates are based on four chains with 1000 samples for adaptation, 3000 samples for burn in and 10000 samples per chain. We assess convergence with the potential scale reduction factor (PSRF, see Brooks and Gelman, 1998).

## B Tables

	Bayesian	Mixed Bayesian
$p(\lambda_2 > \lambda_1)$	100%	100%
$p(\lambda_3 > \lambda_2)$	100%	99.9%
$p(\lambda_4 > \lambda_3)$	42.6%	99.8%
$p(\lambda_5 > \lambda_4)$	75.3%	57.7%
$p(c_1 > c_2)$	99.5%	100%
$p(c_2 > c_3)$	98.9%	83.2%
$p(c_3 > c_4)$	88.3%	84.1%
$p(c_4 > c_5)$	88.0%	74.1%

Table C.2: Attraction effect and precision over time

Note: 111 individuals (63 in graphical treatment, 48 in numeric treatment), 30 menus, 5 time periods



	MLE		Bayesian		Mixed Bayesian		
	mean	conf.int.	mean	conf.int.	mean	conf.int.	st.dev. conf.int.
<b><math>\lambda</math> precision</b>							
t = 4	0.37	[0.11;0.62]	0.38	[0.21;0.56]	1.93	[1.47;2.39]	1.34 [0.97;1.75]
t = 8	1.20	[0.87;1.52]	1.21	[0.86;1.56]	4.66	[3.86;5.47]	2.80 [2.16;3.49]
t = 12	7.91	[7.23;8.59]	7.87	[7.2;8.52]	6.96	[5.95;7.98]	3.87 [2.98;4.74]
t = 16	8.13	[7.36;8.9]	6.33	[5.23;8.41]	9.18	[8.05;10.37]	4.51 [3.52;5.59]
t = 20	8.26	[7.62;8.91]	8.82	[7.84;9.7]	9.35	[8.26;10.48]	4.33 [3.28;5.33]
<b>c competitor malus</b>							
t = 4	1.60	[1.07;2.13]	1.62	[1.19;2.08]	1.23	[1.14;1.32]	0.23 [0.15;0.32]
t = 8	1.12	[1.03;1.2]	1.13	[1.05;1.22]	1.06	[1.03;1.09]	0.10 [0.08;0.13]
t = 12	1.03	[1.02;1.04]	1.03	[1.02;1.04]	1.04	[1.02;1.06]	0.08 [0.07;0.1]
t = 16	1.02	[1.01;1.03]	1.02	[1.01;1.04]	1.02	[1;1.04]	0.08 [0.06;0.09]
t = 20	1.01	[1;1.02]	1.01	[1;1.02]	1.01	[1;1.03]	0.07 [0.06;0.09]
<b>d decoy malus</b>							
t = 4	6.52	[2.56;10.48]	6.65	[4.07;9.61]	2.14	[1.81;2.49]	0.44 [0.18;0.74]
t = 8	2.94	[2.33;3.55]	2.97	[2.37;3.68]	1.60	[1.44;1.8]	0.29 [0.16;0.45]
t = 12	1.19	[1.14;1.24]	1.19	[1.15;1.24]	1.62	[1.41;1.85]	0.30 [0.15;0.46]
t = 16	1.23	[1.18;1.29]	1.36	[1.2;1.48]	1.61	[1.4;1.85]	0.26 [0.14;0.4]
t = 20	1.24	[1.19;1.3]	1.22	[1.16;1.28]	1.55	[1.36;1.76]	0.23 [0.13;0.37]
<b>v value of no choice</b>							
t = 4	-0.40	[-1.37;0.57]	-0.43	[-1.21;0.2]	1.35	[0.92;1.8]	1.81 [1.33;2.35]
t = 8	1.39	[1.25;1.52]	1.22	[1.05;1.49]	1.57	[1.3;1.86]	0.99 [0.71;1.28]
t = 12	-0.84	[-0.96;-0.73]	-2.11	[-5.59;0.29]	1.38	[1;1.78]	0.98 [0.55;1.45]
t = 16	0.70	[0.59;0.81]	1.15	[0.66;1.35]	0.85	[-0.97;2.6]	3.32 [0.62;7.71]
t = 20	1.44	[1.37;1.51]	-1.12	[-5.07;1.55]	1.19	[0.62;1.76]	0.88 [0.27;1.98]
<b>AIC/DIC</b>	33678		33664		26397		

Table C.1: Choice model estimates (MLE, Bayesian, Mixed Bayesian)

Note: 111 individuals (63 in graphical treatment, 48 in numeric treatment), 30 menus, 5 time points.  
95% confidence intervals for ML estimates, 95% credible intervals for Bayesian estimates.

AIC is the Akaike Information Criterion (for ML estimates), DIC is the Deviance Information Criterion (for Bayesian estimates).

	MLE						Bayesian						Mixed Bayesian					
	Numeric		$\Delta$ Graphical		Numeric		$\Delta$ Graphical		Numeric		$\Delta$ Graphical		Numeric		$\Delta$ Graphical		St.Dev.	
	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.
<b><math>\lambda</math> precision</b>																		
t = 4	0.42	[-0.71;1.55]	-0.04	[-1.2;1.13]	0.42	[0.23;0.63]	-0.03	[-0.2;0.15]	2.04	[1.25;2.74]	-0.28	[-1.03;0.53]	1.36	[0.96;1.79]				
t = 8	1.11	[0.37;1.85]	0.38	[-0.51;1.26]	1.09	[0.72;1.46]	0.23	[-0.3;0.75]	5.08	[3.99;6.08]	-0.71	[-1.82;0.47]	2.84	[2.18;3.55]				
t = 12	9.78	[9.78;9.78]	-2.88	[-2.88;-2.87]	9.11	[8.14;10.1]	-2.05	[-3.17;-0.91]	7.45	[6.19;8.74]	-0.91	[-2.23;0.45]	3.94	[3.04;4.82]				
t = 16	10.51	[9.3;11.71]	-4.66	[-6.1;-3.23]	9.16	[8.14;10.17]	-3.03	[-4.18;-1.88]	10.03	[8.61;11.52]	-1.52	[-3.08;-0.12]	4.32	[3.32;5.38]				
t = 20	10.44	[9.33;11.54]	-3.60	[-4.96;-2.24]	9.66	[8.73;10.63]	-2.12	[-3.25;-0.98]	10.40	[9.06;11.76]	-1.93	[-3.38;-0.52]	3.98	[3.01;5.01]				
<b>c competitor malus</b>																		
t = 4	1.54	[0.01;3.08]	0.04	[-1.61;1.7]	1.62	[1.08;2.2]	-0.01	[-0.65;0.64]	1.25	[1.12;1.4]	-0.04	[-0.21;0.15]	0.24	[0.16;0.33]				
t = 8	1.18	[1.1;1.36]	-0.09	[-0.29;0.11]	1.18	[1.03;1.35]	-0.08	[-0.26;0.11]	1.06	[1.01;1.11]	-0.01	[-0.07;0.06]	0.11	[0.08;0.13]				
t = 12	1.04	[1.04;1.04]	-0.01	[-0.01;-0.01]	1.04	[1.02;1.06]	-0.02	[-0.04;0.01]	1.05	[1.02;1.08]	-0.02	[-0.07;0.02]	0.08	[0.07;0.11]				
t = 16	1.03	[1.01;1.04]	-0.02	[-0.04;0.01]	1.03	[1.01;1.04]	-0.02	[-0.04;0.01]	1.03	[1.1;1.06]	-0.02	[-0.06;0.02]	0.08	[0.06;0.09]				
t = 20	1.02	[1.1;1.03]	-0.02	[-0.04;0]	1.02	[1.1;1.03]	-0.02	[-0.04;0]	1.02	[1.1;1.05]	-0.02	[-0.06;0.02]	0.07	[0.06;0.09]				
<b>d decoy malus</b>																		
t = 4	5.79	[-7.35;18.92]	0.49	[-	6.12	[3.7;8.82]	0.42	[-1.34;2.22]	2.16	[1.73;2.65]	0.04	[-0.45;0.51]	0.50	[0.2;0.84]				
t = 8	3.29	[1.6;4.99]	-0.83	[13.39;14.37]	3.39	[2.55;4.39]	-0.66	[-1.74;0.39]	1.51	[1.3;1.72]	0.18	[-0.06;0.43]	0.28	[0.15;0.44]				
t = 12	1.21	[1.21;1.21]	-0.02	[-2.62;0.95]	1.24	[1.16;1.33]	-0.05	[-0.16;0.04]	1.75	[1.43;2.13]	-0.18	[-0.51;0.13]	0.31	[0.15;0.51]				
t = 16	1.22	[1.13;1.31]	0.12	[-0.03;-0.01]	1.29	[1.19;1.4]	0.02	[-0.11;0.15]	1.83	[1.41;2.36]	-0.26	[-0.75;0.16]	0.26	[0.13;0.43]				
t = 20	1.25	[1.15;1.35]	0.02	[-0.01;0.24]	1.30	[1.2;1.41]	-0.07	[-0.2;0.06]	1.57	[1.32;1.84]	-0.02	[-0.28;0.24]	0.24	[0.13;0.37]				
<b>v value of no choice</b>																		
t = 4	-0.78	[-5.46;3.9]	0.93	[-3.85;6.6]	-0.87	[-1.86;-0.05]	0.95	[0.21;1.75]	1.19	[0.51;2.05]	0.40	[-0.49;1.19]	1.90	[1.36;2.5]				
t = 8	1.19	[0.97;1.4]	0.20	[-0.08;0.48]	1.17	[1.02;1.33]	0.26	[0.02;0.52]	1.55	[1.18;1.92]	0.03	[-0.47;0.54]	0.99	[0.71;1.28]				
t = 12	-0.75	[-0.79;-0.72]	-1.94	[-1.97;-1.9]	-2.24	[-6.29;0.42]	-0.24	[-2.13;1.67]	1.27	[0.72;1.85]	0.46	[-0.25;1.2]	1.09	[0.58;1.65]				
t = 16	0.65	[0.56;0.74]	0.54	[0.38;0.7]	0.66	[0.53;0.77]	0.66	[0.36;0.87]	0.50	[-1.66;2.32]	0.51	[-0.9;2.05]	3.44	[0.58;9.97]				
t = 20	1.38	[1.31;1.45]	0.07	[-0.12;0.26]	1.39	[1.31;1.47]	-1.09	[-2.32;0.26]	1.15	[1.04;1.87]	0.12	[-0.86;1.15]	0.97	[0.27;2.14]				
AIC/DIC	33398						33362						26387					

Table C.3: Choice model estimates, by treatment (Mixed Bayesian estimates)

Note: 111 individuals (63 in graphical treatment, 48 in numeric treatment), 30 menus, 5 time points. 95% confidence intervals for ML estimates, 95% credible intervals for Bayesian estimates.

AIC is the Akaike Information Criterion (for ML estimates), DIC is the Deviance Information Criterion (for Bayesian estimates).

		Parameter estimates						Standard deviation				
		$\Delta \text{markup} = +15\%$		$\Delta \text{markup} = +5\%$		$\Delta \text{markup} = -15\%$						
		mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	mean	conf.int.	
<b><math>\lambda</math> precision</b>												
t = 4	1.37	[0.91;1.86]	0.15	[-0.13;0.39]	0.08	[-0.23;0.37]	0.28	[-0.01;0.57]	-0.13	[-0.39;0.16]	1.16	[0.8;1.53]
t = 8	3.25	[2.49;4.08]	1.07	[0.34;1.82]	-0.01	[-0.66;0.66]	0.18	[-0.46;0.84]	0.44	[-0.24;1.09]	1.79	[1.12;2.45]
t = 12	6.06	[4.91;7.33]	0.22	[-0.74;1.15]	-0.26	[-1.11;0.57]	-0.10	[-0.98;0.73]	-0.07	[-0.85;0.77]	3.18	[2.35;4.05]
t = 16	8.96	[7.51;10.41]	-0.42	[-1.5;0.71]	-0.32	[-1.53;0.94]	-1.57	[-2.67;-0.5]	0.77	[-0.49;2.05]	4.08	[3.13;5.14]
t = 20	9.22	[7.79;10.71]	-0.30	[-1.52;0.98]	-0.18	[-1.49;1.13]	-1.35	[-2.56;-0.08]	0.71	[-0.66;2.07]	3.90	[2.89;4.9]
<b>c competitor malus</b>												
t = 4	1.65	[1.38;1.94]	-0.60	[-0.93;-0.28]	-0.45	[-0.75;-0.16]	-0.37	[-0.66;-0.11]	-0.26	[-0.54;0.01]	0.31	[0.19;0.46]
t = 8	1.34	[1.23;1.45]	-0.41	[-0.56;-0.27]	-0.37	[-0.51;-0.24]	-0.31	[-0.43;-0.2]	-0.18	[-0.29;-0.07]	0.13	[0.1;0.17]
t = 12	1.18	[1.12;1.23]	-0.21	[-0.3;-0.13]	-0.18	[-0.25;-0.12]	-0.20	[-0.25;-0.14]	-0.07	[-0.14;-0.01]	0.09	[0.07;0.11]
t = 16	1.10	[1.07;1.14]	-0.12	[-0.18;-0.06]	-0.11	[-0.14;-0.07]	-0.12	[-0.16;-0.08]	-0.08	[-0.12;-0.04]	0.08	[0.07;0.11]
t = 20	1.09	[1.06;1.12]	-0.11	[-0.17;-0.06]	-0.10	[-0.14;-0.06]	-0.11	[-0.14;-0.07]	-0.07	[-0.11;-0.03]	0.08	[0.06;0.09]
<b>d decoy malus</b>												
t = 4	2.44	[1.85;3.08]	0.03	[-0.53;0.58]	0.05	[-0.52;0.63]	0.15	[-0.44;0.74]	0.50	[-0.22;1.2]	0.60	[0.19;1.06]
t = 8	1.86	[1.57;2.16]	-0.24	[-0.57;0.09]	-0.06	[-0.42;0.32]	0.00	[-0.36;0.38]	-0.07	[-0.38;0.26]	0.29	[0.15;0.45]
t = 12	1.63	[1.39;1.88]	-0.02	[-0.31;0.3]	0.20	[-0.16;0.63]	-0.01	[-0.25;0.23]	0.12	[-0.14;0.41]	0.28	[0.15;0.44]
t = 16	1.56	[1.32;1.84]	0.72	[-0.06;1.73]	0.25	[-0.22;0.82]	0.38	[-0.11;0.96]	-0.07	[-0.29;0.14]	0.24	[0.13;0.38]
t = 20	1.47	[1.28;1.68]	0.40	[-0.15;1.15]	0.39	[-0.11;1.07]	0.31	[-0.11;0.85]	-0.06	[-0.25;0.12]	0.20	[0.12;0.3]
<b>v value of no choice</b>												
t = 4	1.68	[1.12;2.3]	-0.06	[-0.28;0.16]	-0.12	[-0.34;0.09]	-0.19	[-0.43;0.04]	0.03	[-0.19;0.25]	2.54	[1.71;3.49]
t = 8	1.60	[1.21;1.99]	0.23	[0.03;0.42]	-0.21	[-0.4;0]	-0.07	[-0.28;0.15]	0.09	[-0.1;0.29]	1.14	[0.81;1.52]
t = 12	1.38	[0.99;1.8]	-0.18	[-0.45;0.08]	-0.02	[-0.27;0.21]	0.12	[-0.13;0.37]	-0.05	[-0.29;0.2]	0.91	[0.53;1.37]
t = 16	1.31	[-1.02;3.58]	-1.27	[-2.55;-0.21]	-0.02	[-0.95;1.02]	-0.23	[-1.54;1.01]	0.16	[-0.72;1.31]	4.38	[0.8;11.4]
t = 20	1.31	[0.17;2.63]	0.73	[-0.08;1.73]	-1.36	[-2.45;-0.45]	0.13	[-0.81;0.91]	1.23	[0.28;2.31]	1.89	[0.49;3.94]

Table C.4: Choice model estimates, by level of the markup (Mixed Bayesian)

Note: 111 individuals (63 in graphical treatment, 48 in numeric treatment), 30 menus, 5 time periods  
95% credible intervals

	Fast then slow (HIM)			Heuristic follower (H)			Maximizer (M)		
	MLE	Bayesian	Mixed Bayesian	MLE	Bayesian	Mixed Bayesian	MLE	Bayesian	Mixed Bayesian
<b><math>\lambda</math> precision</b>									
t = 1	0.39 [0.22,0.56]	0.72 [0.37,1.06]	1.31 [0.88,1.76]	1.32 [1.05,1.59]	1.31 [1.04,1.60]	1.93 [1.52,2.39]	3.35 [2.92,3.78]	5.47 [4.67,6.25]	5.88 [4.74,7.05]
t = 2	1.00 [0.43,1.57]	2.90 [2.28,3.52]	3.32 [2.60,4.04]				8.67 [7.94,9.41]	12.72 [11.14,14.40]	14.37 [12.35,16.68]
t = 3	3.88 [2.99,4.78]	5.42 [4.66,6.19]	6.62 [5.55,7.76]				11.24 [10.35,12.14]	16.44 [14.22,18.63]	18.93 [16.11,21.77]
t = 4	6.30 [5.33,7.27]	6.80 [5.97,7.65]	7.55 [6.38,8.81]				12.26 [11.3,13.22]	19.15 [16.72,21.71]	23.01 [19.28,27.32]
t = 5	7.55 [6.5,8.6]	7.24 [6.39,8.11]	8.13 [6.85,9.39]				12.61 [11.62,13.6]	19.27 [16.85,21.64]	22.14 [19.14,25.74]
<b>c competitor malus</b>									
t = 1	1.23 [0.66,1.81]	1.51 [1.19,1.91]	1.47 [1.22,1.73]	1.12 [1.02,1.21]	1.01 [0.92,1.09]	1.00 [0.94,1.07]			
t = 2	1.23 [1.03,1.44]	1.09 [1.04,1.14]	1.11 [1.05,1.17]						
t = 3	1.03 [1,1.07]	1.04 [1.02,1.06]	1.05 [1.02,1.08]						
t = 4	1.02 [0.99,1.04]	1.02 [1.01,1.04]	1.03 [1.00,1.06]						
t = 5	1.01 [0.99,1.03]	1.01 [1.00,1.03]	1.02 [0.99,1.04]						
<b>d decoy malus</b>									
t = 1	6.46 [4.31,8.62]	3.82 [2.48,5.63]	2.85 [2.12,3.68]	3.19 [2.61,3.76]	3.43 [2.80,4.13]	3.53 [2.90,4.26]			
t = 2	3.33 [1.86,4.81]	1.67 [1.48,1.88]	1.95 [1.65,2.27]						
t = 3	1.52 [1.33,1.71]	1.30 [1.22,1.40]	1.51 [1.33,1.71]						
t = 4	1.29 [1.19,1.4]	1.26 [1.18,1.33]	1.61 [1.38,1.85]						
t = 5	1.23 [1.14,1.31]	1.24 [1.17,1.31]	1.58 [1.36,1.83]						
<b>v value of no choice</b>									
t = 1	-2.2 [-3.67,-0.73]	0.61 [0.30,0.88]	1.13 [0.68,1.59]	-0.57 [-0.92,-0.22]	-2.20 [-2.96,-1.49]	-2.26 [-4.21,-0.45]	1.23 [1.18,1.27]	1.37 [1.31,1.42]	1.36 [1.16,1.57]
t = 2	0.99 [0.77,1.21]	1.42 [1.31,1.54]	1.60 [1.43,1.80]	-0.35 [-0.73,0.03]	-0.33 [-0.66,-0.04]	-0.25 [-0.80,0.29]	1.44 [1.38,1.5]	1.66 [1.47,1.87]	1.98 [1.36,3.30]
t = 3	1.15 [1,1.3]	1.57 [1.47,1.69]	3.62 [1.41,6.85]	-0.03 [-0.41,0.36]	0.34 [0.09,0.58]	0.52 [0.11,0.92]	3.16 [2.26,4.06]	1.09 [0.42,1.76]	3.43 [1.82,7.73]
t = 4	1.38 [1.21,1.54]	-1.87 [-5.45,0.70]	1.08 [4.08,6.73]	0.54 [0.18,0.89]	0.69 [0.37,0.99]	1.19 [0.34,2.12]	2.22 [-319,77,324.2]	0.42 [-5.59,6.73]	1.68 [-4.30,7.35]
t = 5	3.49 [-355.4,362.39]	-1.76 [-5.44,0.87]	-1.18 [-5.27,2.18]	1.64 [1.17,2.11]	2.00 [1.45,2.55]	1.87 [1.38,2.40]	1.51 29022.41,29025.43]	0.30 [-5.66,6.45]	0.63 [-4.87,5.76]
<b>estimated proportion in the population</b>	37% [27%,45%]	55% [45%,65%]	61% [51%,70%]	17% [11%,25%]	21% [14%,29%]	21% [13%,29%]	47% [37%,56%]	24% [16%,32%]	18% [11%,26%]
AIC/DIC	31294	29122	26528						

Table C.5: Choice model estimates (Mixed Bayesian mixture)

Note: 111 individuals (63 in graphical treatment, 48 in numeric treatment), 30 menus, 5 time points. 95% confidence intervals for ML estimates, 95% credible intervals for Bayesian estimates.

AIC is the Akaike Information Criterion (for ML estimates), DIC is the Deviance Information Criterion (for Bayesian estimates).

## C Type assignment

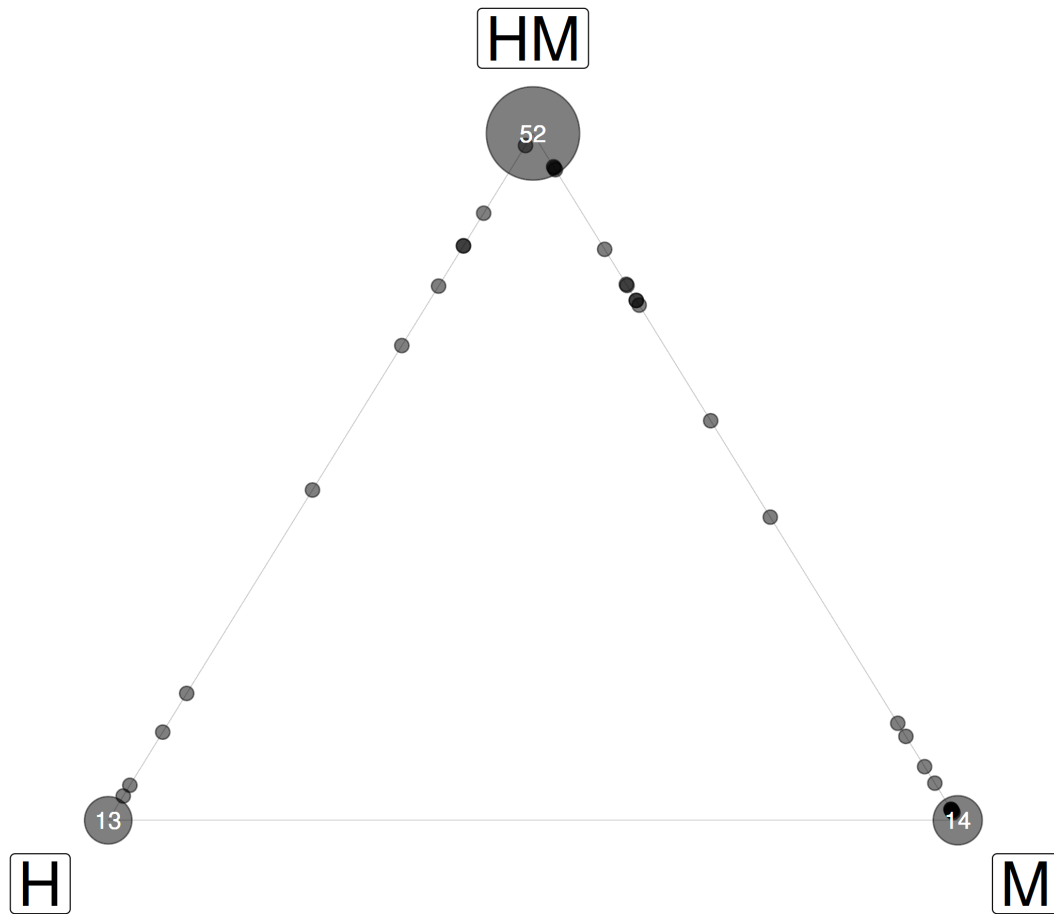


Figure D.1: Type assignment over Heuristic, Maximiser, Heuristic then Maximiser types, and mixtures of the three types.

## D Description of individual statistics

**Difficult:** Was it difficult for you to make choices. Yes=1, No=0

**Experience:** How many studies did you take part in the past?

**Motivation:** Was it important for you to make the right choices? Coded 0-3 in order of importance.

**Problems:** Did you face any problems during this study. Yes=1, No=0

**Understanding:** Did you understand what to do in this study. Yes=1, No=0

**AE:** If two offers were of the same volume and the third was of a different volume, did you tend to prefer or avoid that of a different volume? 0: Prefer competitor, 1: Indifferent 2: Prefer target.

**CRT score:** Sum of correct answers to the standard CRT questionnaire (Frederick, 2005)

**Confusion prone:** Sum of answers to 9 questions adapted from Walsh et al. (2007), each graded from 1 to 5, in order of higher expressed confusion when shopping.

**Loss aversion:** When you decide to take a risk, do you think about the gains you could get, or the losses you could endure? 0-3 in order of tendency to think of losses.

**Risk tolerance:** Do you like to take risks, in general? (Give a number between 0 and 10. 0 if you avoid risks as much as possible, 10 if you always take the maximum risks)

**Trust:** Sum of answers to [Dohmen et al. \(2008\)](#)'s 3 item trust attitude survey, graded higher for higher expressed willingness to trust.

**Male:** 1 if male, 0 if female.

**Age:** in years.

**Revenue:** Monthly revenues, in categories graded as 0=0-1000 €, 1=1000-2000 €, 2=2000-3000 €, etc...

**City size:** Where did you live most of your life, in order of size, from countryside to large city of more than 1 million people.

**Education level:** From no education to advanced tertiary education.

**Economics:** Did you study economics? Yes=1, No=0

**Shopping experience:** Do you generally go shopping yourself? Yes=1, No=0

**Budget holder:** Are you in charge of your own budget? Yes=1, No=0

**Graphical treatment:** =1 if assigned to the graphical treatment.

## E Summary of individual statistics

Variable	N	Mean	St. Dev.	Min	Max
Difficult	111	0.243	0.431	0	1
Experience	111	1.324	2.220	0	20
Motivation	111	2.369	0.660	0	3
Problems	111	0.027	0.163	0	1
Understanding	111	0.982	0.134	0	1
AE	111	1.000	0.556	0	2
CRT score	111	0.919	1.097	0	3
Confusion prone	111	26.459	6.173	11	43
Loss aversion	111	1.459	0.748	0	3
Risk tolerance	111	5.450	2.247	0	10
Trust	111	6.910	1.385	3	12
Male	111	0.378	0.487	0	1
Age	111	26.802	8.232	18	63
Revenue	111	0.360	0.501	0	2
City size	111	2.243	1.046	0	4
Education level	111	3.649	1.211	0	5
Economics	111	0.423	0.496	0	1
Shopping experience	111	0.865	0.343	0	1
Budget holder	111	0.955	0.208	0	1
Graphical treatment	111	0.568	0.498	0	1
Status	Student (59), Worker (36), Unemployed, retired (16)				

Table C.6: Summary of individual statistics

## F Experimental instructions

The original instructions were made up of a Power Point slideshow, using several visual cues as to make them appealing and easy to understand for the varied and heterogeneous population of subjects we faced. The French version of the slideshow is available [ADD LINK](#). Here we provide a translation of all the words, plus the most relevant pictures.

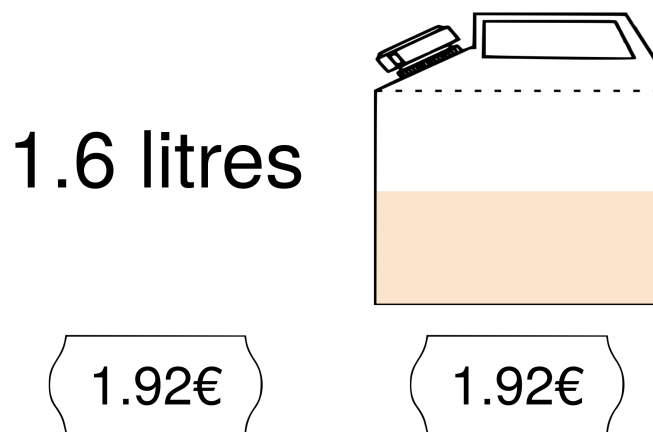


Figure A.1: An offer

## General instructions

Welcome. This experiment is run by the Grenoble Applied Economics Laboratory (GAEL), part of the University Grenoble-Alpes (UGA).

The study is about individual consumption behavior. Instructions will be given to you as we go along. During the whole session, you will have to take some simple decisions. Nonetheless, should you have any difficulty or misunderstandings, do not hesitate to ask.

In order to protect your privacy during the session and in data analysis, you have been assigned a code. No data allowing us to identify you will be collected. Thus, it will be impossible for us to link your replies and decisions to your name. Data will be kept for statistical analysis and publication, but always in their anonymous format.

Communication between participants is not allowed, nor are comments about what should or should not be done during the experiment. Keep concentrated on your own computer screen and keep silence for the whole session. If you have a question, feel free to raise your hands and ask anytime.

You will see on your desk an envelope containing 10 € in cash, rewarding you for your participation. This sum is yours to keep. During the experiment, you will have the opportunity to gain extra money. The amount that you could gain will depend on the decisions you'll make during the session.

This study is composed of a learning phase followed by a payoff-relevant phase. Before the start of each phase, instructions will be given. A new phase will start only after all participants have completed the previous phase. This session will not exceed one hour and thirty minutes.

Do you have any questions?

## Your Task - buying gas

You have to fill with gasoline a jerrycan that can contain 3 liters. You have a budget of 5 € to fulfill the task, and a limited time.

Your gain consists in all the money that you will not have spent on the gasoline. Your task is hence to find the cheapest offer on the market. Unfortunately, you will not have access to the price per liter. Each gas station will show you its offers as a quantity of gas, and the price asked for this quantity.

The quantities are shown [*graphical*: by means of a visual representation] [*numeric*: in liters]

This is an offer. The quantity of gasoline [*graphical*: represented by the pink-colored part] is displayed on top, and the price for the shown quantity below.

With this information, you can make an estimation of the cost of the 3 liters. The price of offers, as

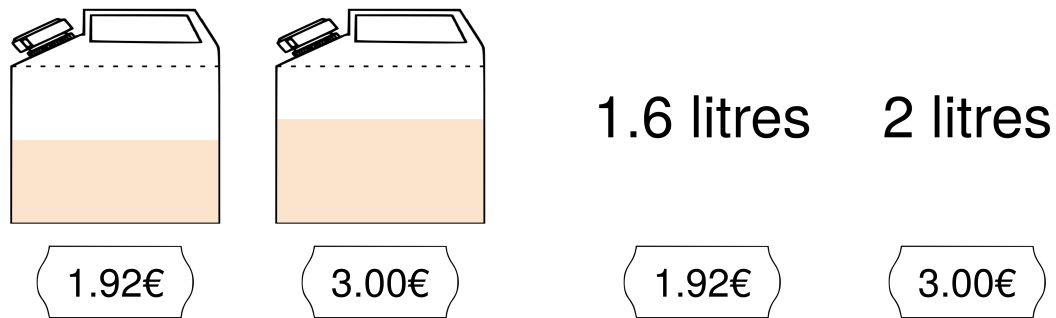


Figure A.2: A two-offer menu

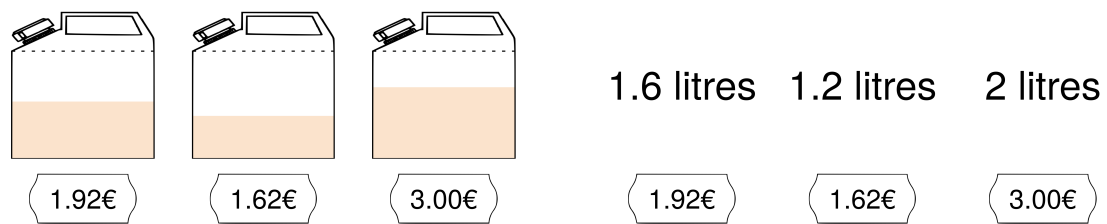


Figure A.3: A three-offer menu

well as the quantities offered for that price, are randomly drawn. The drawings are independent – that is, the prices of the different offers have no particular relationship with one another.

There are two types of gas stations: those displaying two offers and those displaying three. In both cases, the offers are independent from one another. Your task is always the same: finding the least expensive offer.

Once you have chosen an offer, the computer computes automatically the total price that you will have to pay for the three liters. You keep the change: your payoff consists of all the money that you have not spent on the gasoline. To maximise your payoff, you have simply to find the cheapest offer.

Example of payoff computation: if you choose this offer:

- for a price of 1.35 € you will buy 1.5 liters.
- this quantity corresponds, in this example, to half of the target quantity
- your total expenditure will hence be  $2 \times 1.35 = 2.7$  €
- your payoff if you choose this offer is hence of  $5 \text{ €} - 2.7 \text{ €} = 2.3 \text{ €}$

This task will be repeated 40 times, with different offers, in different gas stations with two or three offers. At the end of the experiment, 5 screens will be randomly drawn. Your total payoff will be the sum of the payoffs obtained in each of these 5 screens.

## How do we record your choices

Attention! we will be using a specific process to keep track of your choices.

You will have 20 seconds to make a choice. During these 20 seconds, your choice will be recorded at each moment. You can change your mind as many times as you wish. A bar on the bottom of your screen will inform you on the remaining time.

Now suppose your choice pattern looks like this [graphical: same screenshot with graphical stimuli]:



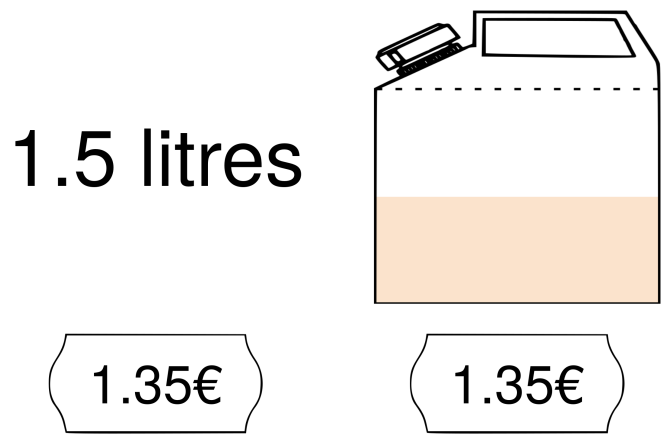


Figure A.4: The offer used in the computation example

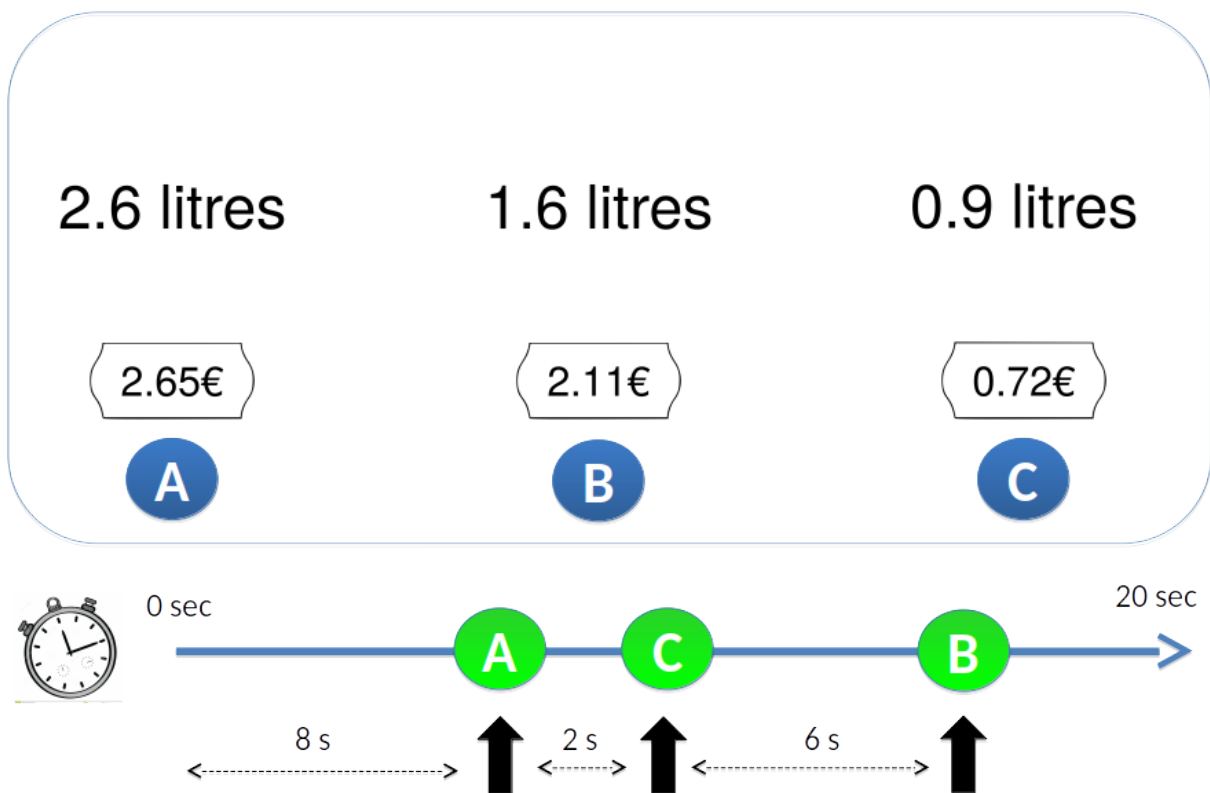


Figure A.5: An example choice pattern

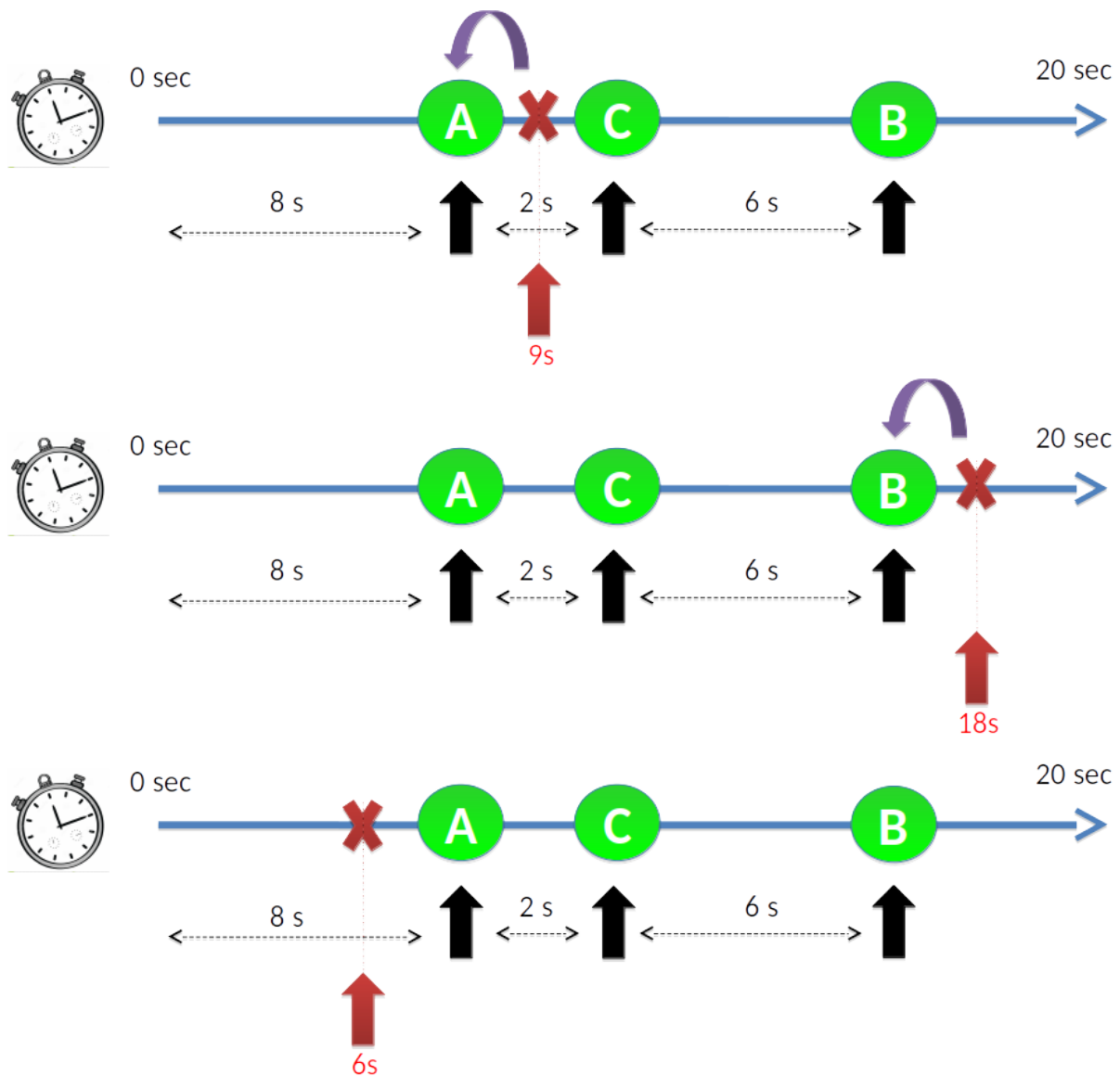


Figure A.6: Three cases for the random draw

At the end of the task, the computer will randomly draw a second between 1 and 20. The payoff-relevant choice will be the one that was active at the randomly drawn time. If no choice had been made at the randomly drawn time, the computer will assign you a random offer.

Let us consider some cases.

**Case1** the time chosen by the computer is 9 seconds. The payoff-relevant choice is A.

**Case2** the time chosen by the computer is 18 seconds. The payoff-relevant choice is B.

**Case3** time chosen by the computer is 6 seconds. No choice! the payoff-relevant offer will be randomly drawn among A, B and C.

Summing up: as long as you made no choice, you risk that the computer selects for you an offer you do not like. It is hence in your best interest to make a fast, possibly provisional, choice, to prevent the random draw from determining your result. You can then change your mind afterwards, as many times as you like.

## **Learning phase**

We will first let you practice on the task. You will face 4 practice screens. At the end of the 4 screens you will be shown a summary of all your choices and potential gains for these screens. This phase is not payoff relevant: the payoffs that you will see on the summary screen are just meant for training purposes.

## **Payoff-relevant phase**

You will now start the payoff-relevant phase. You will face 40 different screens. At the end of the experiment, 5 screens will be randomly drawn. Your total payoff will be the sum of the payoffs obtained in each of these 5 screens.