Do consumers prefer offers that are easy to compare? An experimental investigation

Crosetto, Paolo and Gaudeul, Alexia

8 March 2012

Do consumers prefer offers that are easy to compare? An experimental investigation[*]

Paolo Crosetto[†] and Alexia Gaudeul[‡]

February 8, 2012

Abstract

Consumers make mistakes when facing complex purchasing decision problems but if at least some consumers disregard any offers that is difficult to compare with others then firms will adopt common ways to present their offers and thus make choice easier. We design an original experiment to identify consumers' choice heuristics in the lab. Subjects are asked to choose from menus of offers and we measure the extent to which they favor those offers that are easy to compare with others in the menu. A sufficient number of subjects do so with sufficient intensity for offers presented in common terms to generate higher revenues than offers that are expressed in an idiosyncratic way.

Keywords: Bounded Rationality, Cognitive Limitations, Standards, Consumer Choice, Experimental Economics, Heuristics, Pricing Formats, Spurious Complexity.

JEL Codes: D83, L13, D18

[*] We wish to thank Nicolas Berkowitsch, Alena Otto and Robert Sugden for their comments and suggestions. This paper was presented at the Max Planck Institute for Economics in Jena in April 2011, at the workshop for Experimental Methods and Economic Modeling in Capua in June 2011, at the Max Planck Institute for Human Development in Berlin in June 2011, at the 2011 International Conference of the Economic Science Association in Chicago in July 2011, at the 2011 European Conference of the Economic Science Association in Luxembourg in September 2011 and at the Department of Economics, Business and Statistics (DEAS) of the University of Milan in December 2011. The experiment's interface and randomized menu generation were programmed with Python (van Rossum, 1995). Data analysis and regressions were performed with Stata (StataCorp, 2009). Simulations were run with Octave (Eaton, 2002).

[†] Strategic Interaction Group (ESI), Max Planck Institute for Economics, Jena. email: crosetto@econ.mpg.de

[‡] Graduate School "Human Behavior in Social and Economic Change" (GSBC), Friedrich Schiller University, Jena. email: a.gaudeul@uni-jena.de

Behavioral economics finds that consumers have "inconsistent, context dependent preferences" and may not have "enough brainpower to evaluate and compare complicated products" (Spiegler, 2011). They "may fail to choose in accordance with what, after sufficient reflection, they would acknowledge to be their own best interests" (Gaudeul and Sugden, 2011). Low levels of consumer literacy and numeracy even in advanced economies make it very difficult for broad swathes of the population to understand how to make adequate decisions in many situations, such as when choosing how much to save for retirement, when selecting healthcare insurance, when investing in stock markets, when comparing car or computer models, *etc.* (Agarwal and Mazumder, 2010; Ayal, 2011; Bar-Gill and Stone, 2009; Lusardi, 2008; Miravete, 2003; Wilson and Price, 2010).

Marketing research (Morwitz et al., 1998; Nunes, 2000; Viswanathan et al., 2005; Zeithaml, 1982) and research from behavioral economics (Ariely, 2008; Iyengar and Lepper, 2000; Iyengar et al., 2004) gives examples of how badly consumers deal with products choices in realistic purchasing scenarios. Experiments on this topic include Huck and Wallace (2010), Choi et al. (2010) and Shestakova (2011) among others.

The consumers biases, limitations and inconsistencies that are evidenced in such research have consequences in terms of strategy for firms (Ellison, 2006; Spiegler, 2011). Firms may benefit from introducing spurious complexity in their contract offerings so as to deliberately obfuscate consumer choice (Carlin, 2009; Chioveanu and Zhou, 2009; Ellison, 2005; Gabaix and Laibson, 2006; Piccione and Spiegler, 2010). To use a term introduced by Adams (1997), sectors in which firms do so are "confusopolies". This is defined as "a group of companies with similar products who intentionally confuse consumers instead of competing on price". Sectors in which this might be the case include telephone services, insurance, mortgage loans, banking, financial services, electricity, *etc.* In all those sectors, firms sell a relatively homogeneous product and so would make low profits if they did not introduce spurious differentiation in their offerings and thus undermine consumers' ability to make informed choices about their services and products. Recent research does find empirical evidence that firms might design their offers to exploit consumers (DellaVigna and Malmendier, 2006; Ellison and Ellison, 2009; Miravete, 2003, 2011). Kalaycı and Potters (2011) also found experimental evidence that more complex offers increase firms profits in a duopoly setting

Faced with such issues, libertarian paternalists (Camerer et al., 2003; Thaler and Sunstein, 2008) suggest regulatory intervention to impose that consumers' decision problems be framed in such a way that they reach the "correct" decision, that is, the decision they would take *absent their limitations.* However, determining what decision that would be is difficult, not to mention that even experts may not know what is best (Freedman, 2010). A complementary option is to introduce measures to educate consumers and provide them with information so they have the tools to make better choices in a wide variety of settings (Agarwal et al., 2010; Garrod et al., 2008).

Before even considering such interventions, one has to prove they are needed, and Sugden (2004) argues that they are not. He maintains that consumers ought to be left free to choose as they wish and the market left free to fulfill their needs as they occur. Sustaining this argument, Gaudeul and Sugden (2011) show that competition will drive firms to simplify their offerings on their own if at least some consumers discard offers that are difficult to compare with others. This is what they call the common standard effect. The common standard rule (choose only among offers that are easy to compare with others) is a rule of thumb that assists consumers in their selection of which product to buy. An example of how it operates goes as follows: a consumer wants to buy a fruit and is faced with the choice between two oranges and one apple. Oranges are priced at \$0.45 and \$0.55 respectively, while the price of the apple is \$0.70. Suppose the consumer cares only about calories and estimates the oranges to contain 35 calories each while he thinks the the apple contains 55 calories. The consumer discards the higher priced orange from his consideration set and compares the lower priced orange with the apple in terms of price per calories. From the price and calorific content of each fruit, he calculates that the lower priced orange costs \$1.29 per 100 calories, while the apple costs \$1.27 per 100 calories. The lower priced orange appears to cost more than the apple, but the consumer still chooses it under the CS rule. We will see this makes sense as long as the consumer is not sure about how different fruits compare in terms of calorific content (he knows he might have made mistakes in his evaluation), there is little intrinsic differences between products (he cares only about calories), and the consumer does not hold prior beliefs on the value of each product (he does not believes for example that apples are always the best deal). This rule derives strength from its simplicity, has strong behavioral foundations and can be applied in many settings, thus ensuring its evolutionary robustness. Contributing to the later, we will see that there is no need for others to follow it for it to be optimal.

To clarify our meanings, what we call a "standard" here is what others have called a "frame", that is, to paraphrase Spiegler (2011, p.151), an aspect of a product's presentation that is of no relevance to a consumer's utility and yet affects his ability to make comparisons among alternatives. This can be a price format, the language in a contract clause, but also a unit of measurement, a way of packaging a product, a technical standard, *etc.*... Expressing an offer in terms of a common standard does *not* inherently make that offer less complex to understand. That is, a CS offer when standing on its own will not be easier to evaluate than an offer that is presented in terms of an individuated standard ("IS"). It is only when put in relation with other offers that a CS offer will be easier to evaluate than an IS offer. To take an example, the switch by Apple from PowerPC processors to Intel x86 processors in 2006 did not make the performance of Apple computers easier to evaluate, but it did make it easier to compare with the performance of most other computers. Our argument is thus not an argument about *complexity*, but about *comparability*.

This paper contributes to the experimental literature on consumer decisions in complex settings by exploring whether consumers make use of common standard information, and if so, how much and in which way. We identify what choice heuristics consumers follow when they face menus of offers that are individually difficult to evaluate but can be compared with others. We focus in particular on whether consumers favor offers that are easier to compare.

Subjects are asked in our experiment to buy paint to cover a given fixed area, known and the same for all. They are presented with menus of offers, whereby an offer is presented in terms of the price of the offer and the area the offer can cover for that price. The areas are presented in different shapes (circles, triangles and squares) of different sizes, all smaller than the total fixed area to be covered. It is therefore not only difficult for our subjects to assess how much each offer would cost to cover the given fixed area, but also difficult to determine which of the offers in a menu is the least expensive. However, offers that are of the same shape and size (of the same "standard") are easy to compare, the lower priced of the two being unambiguously a better deal.

We find that our subjects generally obtain better payoffs when a menu includes some offers that are expressed in terms of a common standard, that is, when some offers within a menu are easy to compare. We also observe that a number of consumers favor the lower priced of the CS offers ("LPCS"). While only a minority of consumers follow this "common standard rule", the intensity of their preference for the LPCS ensures that products expressed in terms of a common standard generate higher revenues than others.

We chose to rely on a controlled laboratory setting because empirical data is not well suited for our purpose. Relying on product sales, for example, introduces various confounds: the presence of real along with spurious product differentiation; regulations that impose standards for a variety of reasons; economies of scale and network effects that encourage the convergence to a technological standard; reputation concerns that lead firms not to wish to confuse consumers; framing other than the standard adopted by the offer that may influence choice as well; habits such that the consumer chooses a product based on past purchasing behavior, and so on. Doing an experiment in the laboratory allows us to create genuine spurious complexity, that is, complexity that all consumers would agree should be irrelevant to their choice. We kept the laboratory experience close to a purchasing act by framing the experiment as a real buying decision in which the participants were asked to *buy* a product out of *menus* of *offers* with the aim of minimizing expenditure. This means that even though the task was cognitively complex and making correct choices was difficult, our subjects were still able to easily understand the task they were asked to perform.

1 Experimental design

Our subjects were first faced with a purchasing tasks, which constitutes the core of our experiment, and then had to complete a set of control tasks and fill out a questionnaire. The next section describes the main task.

1.1 The main task

In order to explore consumer behavior when faced with a problem that is both simple to understand but complex to solve, we designed a novel purchasing task with a simple structure in which complexity was introduced in a natural way. Subjects were given a budget B to buy *gray paint* in order to cover a fixed, square area A. They were presented with menus consisting of a number of offers, each offer being expressed in terms of its price and a visual representation of the area that the paint could cover for that price. Formally, each offer was a triple (s, a, p) in which s is a shape, a is the area of the shape s, expressed as a fraction of the total area A, and p is the price of the offer. Participants were told that paint quality did not differ across offers. The subjects' payoff was what remained from their budget B once all the paint needed to cover A had been bought at the cost implied by the chosen offer. The overall price paid for the chosen offer was calculated as p/a, and the payoff for the subject was $B - p/a$.

While the task is conceptually very simple and relates to everyday activities - subjects must minimize expenditure when buying a product of standardized quality - it is also cognitively quite hard, as evaluating hidden unit prices and comparing areas of different shapes can be difficult. Presenting offers in terms of a combination of a shape and a size allowed us to introduce a relatively high level of spurious complexity in an intuitive way while drawing on an existing body of research on shape perceptions (Krider et al., 2001). The concept of a *standard* was also easily introduced within our design: two offers within a menu that shared the same shape *and* size were easy to compare in terms of price, as price was then the only remaining differentiating factor. We therefore denote in our setting an offer as being a *common standard* offer if it has an equivalent in terms of size and shape in the menu.

Since the basic task (choosing an offer within a menu) was repeated several times, we wanted to exclude by design the possibility for our subjects to learn some specific pattern in the offers. Our offers could thus take three different shapes, each of twelve possible sizes, meaning that there were 36 possible distinct standards. Prices themselves were randomly generated, meaning that it was almost impossible for consumers to rely on past purchasing experiences within our experiment to inform their present purchasing task.

The offers' three dimensions varied in the following way:

1. The shape s could be a circle, a square, or an equilateral triangle. We considered only those three shapes so as to be able to build on the existing literature on shape comparisons (Krider et al., 2001).

2. The area a took one of 12 possible values. Normalizing A to 100, these values ranged from 10 to 43, in steps of 3.[1] The step was chosen to be big enough to allow our subjects to determine easily whether an offer was bigger than another of the same shape within a menu, while being small enough to yield a sufficient number of steps and therefore a sufficient number of different (s, a) pairs in order to minimize learning from comparisons across menus.

3. The price information conveyed to the subjects, p, was computed from randomly drawn *unit prices* (up, the cost to cover 1% of A) as $p = up \cdot a$. Unit prices were drawn from a normal distribution of mean 0.5, while standard deviation σ^2 was equal to either 0.05, which generated more distance between offers and hence an *easier* problem, or 0.01, which generated closer offers and thus made it *harder* to identify the best one.

The offers were displayed as a gray area centered on a white background representing the total area to be painted. The triangular offers rested on their base while square offers rested on a side. The white background allowed participants to visually appreciate the size of the shape with respect the total area to be painted. This background was overlaid with a grid of thin light blue lines to ease comparison between offers of the same shape. This made it possible for participant to assess if two offers of the same shape were indeed of the same size.

The offers were displayed in menus, that varied in length (*3* or *6* offers per menu). Menus were randomly generated under the constraint that no offer was to give a negative payoff to the participant. With respect to CS, menus could feature *no* common standard, such that a given (s, a) combination would appear only once within the menu; *one* common standard, such that two (and only two) offers featured the same (s, a) combination; or *two* common standards (only possible for menus of six offers), whereby one (s, a) combination occurred twice while another occurred thrice.

An example of a menu with three elements and a common standard (the triangle) is shown in figure 1. An example of a menu with six elements and no common standard is shown in figure 2.

[1]The size was limited to 43 as an equilateral triangle resting on a base cannot cover more than $(5\times\sqrt{75})/100 = 43.3\%$ of a 10×10 square.

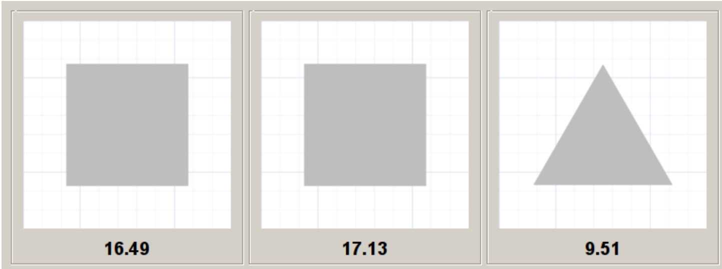

Figure 1: Screen shot of a menu with three offers and a common standard

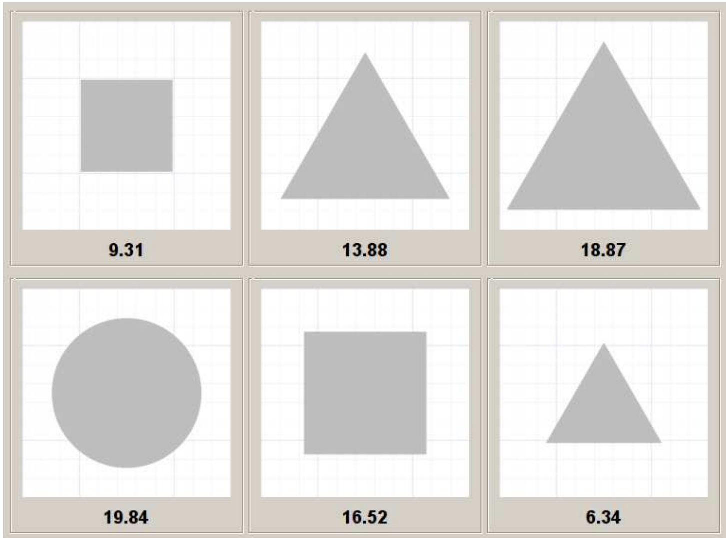

Figure 2: Screen shot of a menu with six offers and no common standard

Each individual was faced with 80 menus, the same set for everyone but presented in a subject-specific random order. 36 menus showed three options ("3-menus"), of which 18 with one CS. 44 showed six options ("6-menus"), of which 18 with one CS and 8 with two CS (one CS with two members, the other CS with three). In each case, half of the menus were *hard* ($\sigma^2 = 0.01$) while the other half were *easy* ($\sigma^2 = 0.05$). The distribution of menus is summarized in table 1 .[2]

Table 1: Distribution of menus by CS and difficulty of the problem

		Hard menus ($\sigma^2 = 0.01$)	Easy menus ($\sigma^2 = 0.05$)
3-menu	No CS	9	9
	One CS	9	9
6-menu	No CS	9	9
	One CS	9	9
	Two CS	4	4

Given the random process governing unit price generation, the lowest priced common standard offer had a theoretical chance of being the optimal choice in $\frac{2}{3}$ of our 3-menus with a CS, in $\frac{1}{3}$ of our

[2]The menus are available for visual inspection at https://people.econ.mpg.de/~crosetto/Shapes/Menus.html.

6-menus with one CS, and in $\frac{5}{6}$ of our 6-menus with two CS. The actual realization of these chances was 0.55 in 3-menus, $\frac{1}{3}$ in 6-menus with one CS and 0.875 in 6-menus with 2 CS.

The participants had up to two minutes to choose an offer from each menu and were forced to spend a minimum time of 10 seconds on each menu. The choice was performed by clicking on an offer - in which case it would be highlighted with a light green frame - and could be revised as many times as one wanted within the two minutes limit. The choice was finalized by clicking on a 'Submit' button at the bottom of the screen. If no final choice was submitted within the time limit the last highlighted offer was submitted as the final choice; if no offer had been highlighted, then the participant received a payment of 3 euros for that trial, which was less than the minimum payment a participant could get even if he made the worst choice out of all our menus.[3]

The participants were given feedback after each menu. This feedback reminded them of the price of their chosen offer, told them the resulting expenditure to paint A, as well as their payoff in terms of budget minus expenditure. The participants were not given the possibility to automatically store and retrieve their payoffs from previous rounds, but were provided with pencil and paper and some did record their payoffs. After the feedback dialog, they were given a new budget B and shown the next menu. The participants knew the total number of menus was 80 and were reminded of their progress along the experiment.

1.2 Control tasks

Once finished with the main task, the participants were exposed to a set of non-incentivized visual perception and computational skills tasks to control for their ability to perform the main task. No minimum time was enforced and the participants could skip any question within each task.[4] Three different set of tasks were chosen:

1. *Shape size comparisons*: The participants were asked to give their estimate of the relative size of a shape (rectangles, circles and triangles) with respect to another. Each of four comparison had to be done within a time limit of one minute.

2. *Mathematical operations*. The participants were asked to solve three sets of 10 operations (sum, subtraction, multiplication, divisions).[5] Each set had to be completed within one minute.

3. *Simple problems*: The participants were asked to solve four simple problems, testing their understanding of the concept of area, of how an area relates to its dimensions, and how a number can be translated from one standard to another (here, a currency). Each problem had to be solved within two minutes.

Once done with the control tasks, the participants filled in a short demographic questionnaire. They were finally asked to guess what the experiment was about - to check for demand effects - and to rate their level of motivation during the experiment. Finally, each participant individually drew a number

[3]Only one participant failed to submit a decision within the time limit, and this only once, in that case highlighting no offer.

[4]Only one participant did so.

[5]The sets were generated using Mail Goggles's GMail Labs app by Jon Perlow and were graded in terms of difficulty. See http://gmailblog.blogspot.com/2008/10/new-in-labs-stop-sending-mail-you-later.html

from 1 to 80 from an urn and was paid according to the result of her purchasing decision in the period corresponding to that number.

Our whole experiment was computerized. The experimental software, the menu generator and the script we used to collect and organize the raw data were programmed in Python (van Rossum, 1995).[6] The German instructions, as well as their English translation, are available upon request.

2 How can (or should) consumers make choices

There are many ways in which one may model consumer choice among offers in our menus, but we will limit ourselves to combinations of two simple criteria for choosing between products: based on imperfect observation of unit prices (what we will call "signals"), and based on whether the product belongs to a CS or not. Other choice factors will be evoked in part 2.4.

Denote $u\hat{p}_{ij} = up_i + e_{ij}$ the perceived unit price of offer i by consumer j. up_i is the unit price of offer i, while e_{ij} is an error term, which is independent across offers in a menu and across consumers. How large the error term will be on average will depend on the consumer's accuracy and on how difficult it is to compare offers *across* standards. As for whether an offer belongs to a CS or not, this matters because prices are directly comparable *within* a standard, so the consumer can identify the LPCS with high accuracy.[7] From those two criteria, we can derive four possible heuristics, illustrated in the following graph and explained below.

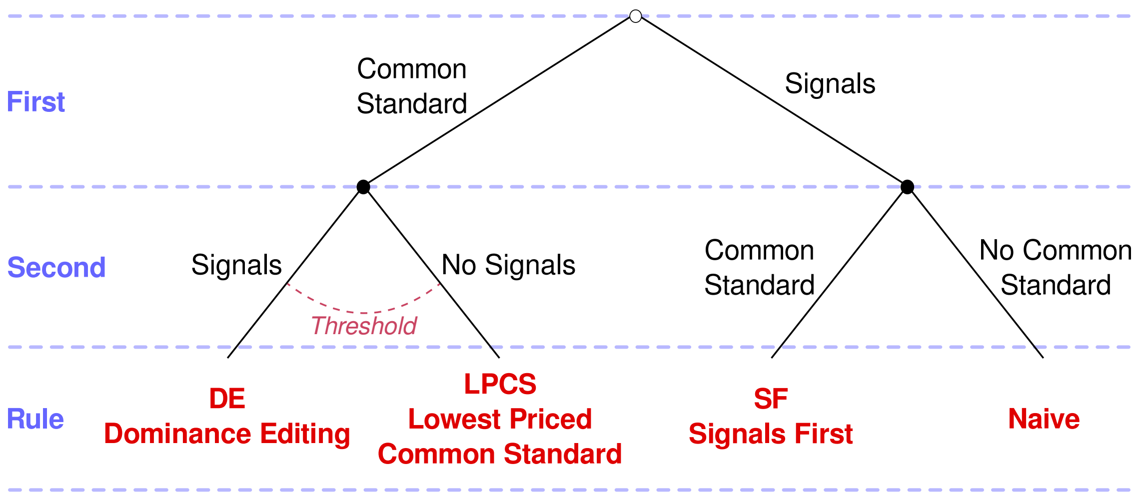

Figure 3: Choice criteria and heuristics

On the left, if the consumer first considers whether offers belong to a CS, he will then eliminate all higher priced CS offers ("HPCS"). From this point on, he may end his search by choosing the LPCS (this is the *CS rule*), or he may compare the signal of the LPCS with that of the individuated standard offers ("IS") in the menu and choose the offer with the lowest signal, which is what we call *Dominance*

[6] Different python modules were needed to develop the experimental software: wxpython was used for the graphical user interface, and two community-contributed packages, svgfig and polygon, were used for creating and managing the shapes. The experimental software (menu and shape generators and analyzers, user interface) and its documentation, as well as the raw data and the script used to collect and organize them are available upon request.

[7] We will consider the possibility that a consumer may make mistakes in choosing among CS even if he is aware of their existence, though one may alternatively argue that choosing a higher priced CS offer means the consumer does not take account of CS information.

Editing ("DE"). On the right, if the consumer first considers signals, he may provisionally choose the offer with the lowest signal. If he does not take account of the existence of a common standard, then he will opt for that offer, thus following what we call the *Naive* rule ("Na" rule). If on the other hand the consumer takes account of the existence of a CS and the offer he provisionally chose turns out to belong to a CS, he will check whether his provisional choice is the LPCS, and if not, revise his choice and opt for the LPCS. This is what we call the *Signal-First* rule ("SF" rule).

In other terms, the Naive rule chooses $\arg\min_i \hat{up}_{ij}$, the CS rule chooses $\arg\min_{i \in CS} p_i$ (the LPCS) if a CS exists and reverts to the Naive rule otherwise, the DE rule determines $k = \arg\min_{i \in CS} p_i$ if there is a CS, in which case it chooses $\arg\min_{i \notin CS} (\hat{up}_{kj}, \hat{up}_{ij})$ and reverts to the Naive rule otherwise, the SF rule determines $l = \arg\min_i \hat{up}_{ij}$ and then chooses $\arg\min_{i \in CS} p_i$ if $l \in$ CS, otherwise chooses l.

2.1 The CS rule

We concentrate in this paper on the CS rule, which is such that consumers not only avoid the higher priced of the common standard offers but choose the lower priced of the CS offers (the "LPCS") and disregard individuated standard ("IS") offers. There are many reasons why we would expect consumers to follow such a rule:

1. *Statistically*, if one assumes that prices are i.i.d. across offers and offers are assigned to a CS at random, then the LPCS is lower priced in expectation than other offers. As in the Monty-Hall problem (Friedman, 1998), there is information gained from being told that an option is dominated.

2. *Behaviorally*, consumers have been shown to be subject to the asymmetric dominance effect (Ariely, 2008, Chapter 1), so that when faced with three offers, one being dominated by another, that other will be chosen more often than if the dominated offer was not present. Another way to call this effect in the field of decision theory is the "attraction effect", which is a type of context effect (Huber and Puto, 1983).

3. *From learning*: Gaudeul and Sugden (2007) argue that consumers are better off choosing among CS offers when firms are strategic agents in a competitive setting, subject to at least some agents following the CS rule. This learning is made easier by the applicability of the common standard rule to many environments, so that consumers who learned from one environments that CS offers are lower priced than other offers will apply this insight generally. Consumers ought therefore to learn to choose CS offers over time (Sugden, 1986; Fudenberg and Levine, 1998).

4. *For simplicity*, as agents faced with complex choices tend to follow simple heuristics, often with good results (Gigerenzer and Brighton, 2009). In this case, an offer being unambiguously better than another provides "one good reason" to choose it (Gigerenzer and Goldstein, 1999).

The CS rule, based on multiple foundations, can thus be generalized across many settings and is likely to be more robust than rules that hold only in some settings (Sugden, 1989) or that can be justified in only one way. We believe this rule is at work in a wide variety of consumer choice problems.

Its simplicity and intuitive appeal make it particularly interesting for economists interested in consumer behavior and heuristics, marketing, consumer protection and the competitive process. Note that we are not wedded to one particular explanation for why consumers might prefer CS offers: we are only interested in determining if they do so and if so, to what extent. Indeed, the main reason we are interested in this possible consumer bias is that we believe that it could drive firms into making their offers less difficult to compare and thus encourage the efficient working of competitive markets. Our setting provides a lower bound for the CS effect, in so far as any competitive effect justifying the use of the rule is excluded by design since offers are not determined through a competitive process.

2.2 The Threshold rule

We will see later that the CS rule is not amenable to econometric analysis as its predictions are too sharp. This is why we will consider the more general *Threshold rule*, of which the DE and LPCS rule are extreme cases. The Threshold rule function as follows: choose $k = \arg\min_{i \in CS} p_i$ if there is a CS and then choose $l(v_j) = \arg\min_{i \notin CS} (\hat{up}_k, \hat{up}_i v_j)$, with threshold v_j depending on consumer j's preference for ($v_j > 1$) or against ($v_j < 1$) the LPCS. The optimal choice of threshold v_j is $v_j^* = \arg\min_{v_j} E(up_{l(v_j)})$. Its level depends on the consumer's accuracy in assessing the unit price of offers in a menu, with less accurate consumers benefiting from adopting higher thresholds v_j. Threshold $v_j = 1$ corresponds to the DE rule, while threshold $v_j \to \infty$ corresponds to the CS rule.

To put this in behavioral terms, the consumer who adopts a threshold $v_j > 1$ does not reject IS offers out of hand, but penalizes them, that is, he does not follow his first impression ($\hat{up}_{ij}$) of the value of the product, but rather revises it upwards when comparing it to his perception of the value of common standard offers. In other terms, the consumer applies a certain dose of skepticism to his evaluation of an offer that is expressed in uncommon terms, and will choose to buy it only if it seems sufficiently better than the best of those offers that are expressed in common terms – that is, its unit price appears to be lower by a factor of at least $1 - 1/v_j$ compared to the apparent unit price of the LPCS.

To make this clearer, let us come back to the example on page 3. We saw that under the CS rule, the consumer would always choose the orange. Under the threshold rule, the consumer will choose the orange only if his threshold v is more than $1.29/1.27 = 1.016$. Section 2.3 on the next page shows that a consumer's threshold ought to depend on his accuracy in assessing offers, with less accurate consumer being better off adopting higher thresholds.

Note that following the CS rule is strictly optimal in the context of Gaudeul and Sugden (2011) as IS offers are systematically higher priced than CS offers in a competitive setting where firms can choose their standard, so that even an IS offer with a very good signal should be rejected. However, *the CS rule is not optimal in the context of our experiment* as offers are randomly generated rather than the result of a competitive process. It is therefore always better for a consumer to follow the Threshold rule with threshold $v_j > 1$ but not infinite as in the CS rule.

We will see later on that no consumer followed the CS rule in our experiment, but a number of them did follow the Threshold rule. The next section goes further into comparing the performance of the various rules introduced above.

2.3 How do the different rules perform?

How the different rules perform depends on how accurate consumers are in their choices. There are two extreme cases: If consumers make no mistakes, then the Naive rule works best and the CS rule is the worst. Indeed, consider for example a 3-menus. Denote B the consumer's budget and a, b, c i.i.d. random variables. The perfectly accurate Naive consumer's expected payoff is $B - E(\min(a, b, c))$, which is more than $B - E(\min(a, b))$, his expected payoff if he restricted himself to CS offers. On the other hand, a consumer who makes considerable mistakes obtains $B - E(a)$ in expectation under the Naive rule (he chooses essentially at random), which is less than $B - E(\min(a, b))$, his expected payoff under the CS rule.

We performed simulations with Octave (Eaton, 2002) to examine the performance of each rule in terms of expected consumer payoff.[8] We modeled e_{ij} as following a normal distribution with mean zero and variance σ^2. In the same way as in our experiment, products unit prices up_i followed a normal distribution with mean 0.5 and variance 0.01 (hard menus), and 0.05 (easy menus) and B was set to 60. Consumer choice was simulated according to the various rules expressed above (Naive, DE, CS, SF), as well as according to the Threshold rule ("Th"), with the optimal threshold v calculated for every level of σ^2 since less accurate consumers benefit from adopting higher thresholds. Their average payoff for each rule was calculated over 2 million menu draws so as to achieve good accuracy.[9]

The following graphs show payoffs in the four situations in our experimental setting, that is depending on whether the consumer has a choice among three or six options, and whether menus are easy or hard. Also shown on separate scale is the optimal threshold v^* for each value of the error term.

[8] Program available upon request.

[9] The ranking of payoffs by rules is quite robust as differences in payoffs are significant even for much smaller draws.

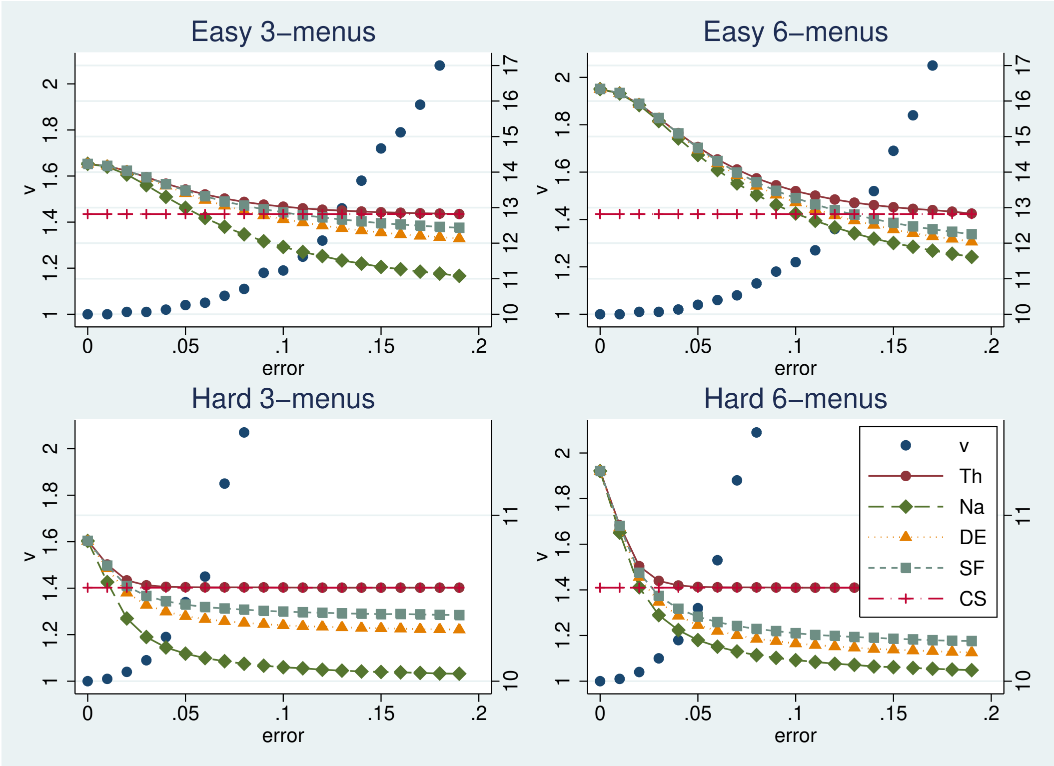

Figure 4: Consumer payoffs by choice rules and optimal thresholds, by menu length and difficulty

As can be seen in figure 4, payoff decreases as consumers become less accurate in their choice (higher levels in σ^2), except for the CS rule since consumers always choose correctly among CS offers and thus obtain $B - E(\min(a, b))$. The Threshold rule outperforms all other rules, and converges towards the CS rule for less accurate consumers. Following the CS rule obtains higher payoffs than the DE, SF or Naive rules as long as consumers are not too accurate. The CS rule is better than those other rules even for rather precise consumers and especially when menus are hard, as even high levels of accuracy may result in mistakes if prices are close together. In terms of ranking, the Threshold rule outperforms the CS rule, while both SF and DE dominate the Naive rule, which is because they take account of the existence of a CS. The reason SF dominates DE is that DE does not recognize that the LPCS is statistically of higher expected value than IS offers, while the SF does not have such a bias against CS offers, treating them in the same way as IS offers in its first step. However, from a practical point of view, DE saves time and effort compared to SF because it requires estimating the value of a lower number of alternatives.

2.4 Other possible rules

Consumers may follow other rules than the ones studied above, and we controlled for those as well. We label those alternative possible rules as follows:

- The *budget rule* chooses arg min p_i. This is a rule that favors small packages, or equivalently lower priced items. While this does not make sense in our setting, this rule may be imported from other settings where for example the consumer faces a binding budget constraints Viswanathan et al. (2005). Alternatively, a *bulk purchasing rule* would favor big packages, as offers in big packages are usually better deals than those in small packages.

- The *lexicographic rule* may favor the first offers in the lexicographic order in the menu – maybe because the consumer is satisficing rather than optimizing (Simon, 1955) or simply because he does not have time to consider all offers. Alternatively, a consumer may also favor the last offers in the menu if he tends to remember (and choose) the last option he read from a list.

- The *shape rule* may favor some shapes over others, as evidenced in Krider et al. (2001).

3 Descriptive statistics and exploratory data analysis

Our experiment took place at the laboratory of the Max Planck Institute in Jena in June 2011. The experiment involved 202 students over 8 sessions, each with 24 to 27 subjects. Our subjects were asked for their age, gender, field of study, year of study, motivation in completing the tasks, and also what they thought the experiment was about (in order to control for demand effects). All subjects were students. When asked what they thought the experiment was about after going through it, most subjects guessed we wanted to assess their abilities to take account of both price and area to identify the best offer in our menus. Some wondered if we wanted to identify what shapes were perceived as more attractive, but no subject mentioned that some offers were expressed in terms of a common standard.

Table 2: Summary statistics

Variable	Mean	Median	Std Dev.	Skewness	Min	Max	N
Age	23.65	23.00	3.69	2.31	18.00	47.00	202
Gender	0.65	1.00	0.48	-0.64	0.00	1.00	202
Score in shape comparisons	0.25	0.25	0.10	0.35	0.05	0.58	201
Score in simple problems	2.78	3.00	0.96	-0.27	1.00	4.00	202
Score in mathematical tasks	20.92	21.50	2.93	-1.45	6.00	25.00	202
Reported motivation	6.29	7.00	2.28	-0.67	0.00	10.00	202
Payoff	11.44	11.48	0.41	-0.80	9.88	12.28	202
Time spent per menu	19.67	18.34	6.36	1.30	11.66	46.27	202

The average age of our subjects was 24, ranging from 18 to 47 (Table 2). 65% of our subjects were women. The average motivation of our subjects, on a scale from 0 to 10, was 6, with a median motivation of 7 and 75% of our subjects having motivation more than 5, the middle point. The monotony of the tasks did not therefore result in noticeable discontent. Speed of choice for each menu and each subject was also recorded. Subjects took 20 seconds on average to make each choice (they could not make a choice before 10 seconds had elapsed). Time spent on each menu was longer for menus with more options and declined over time (from an average of 36 seconds for the first choice to 16 for the last).

There were three control tasks. In the shape comparison task, subjects were asked to assess the area of one shape in terms of multiples of another. We computed individual performance as the average of |guess – true value|/true value. On average, people were 25% off the true value, with a minimum of 5% and a maximum of 58%. In the mathematical tasks, we coded answers as either right or wrong. On average, subjects got 21 of the 25 calculations right, with only two obtaining less than half of the calculations right, and 7 of them obtaining all of them right. Finally, in the simple problems, only about 62% answered more than half of the questions correctly. Performance in the different control tasks were significantly and positively correlated, though not highly (correlation coefficients were around 0.35). Women performed less well than men in all control tasks.

Individual choices, payoffs and performance Overall, consumers made about 39% of their choices correctly, that is, choosing the offer with the lowest unit price. In only 21 of the 80 menus did a majority of the consumers make the correct choice. In other terms, most consumers were wrong for most menus.[10] Table 3 shows that the LPCS was chosen about 57% of the time within our 3-menus,[11] about as often as the LPCS was the lower priced product (56%). This was less often than if consumers followed the CS rule, whereby the LPCS would always be chosen. However, the IS was disfavored as it was chosen less often than if consumers always chose the lowest priced product (37% of the time vs. 44% if choice were optimal). In the case of 6-menus with one CS, the LPCS was chosen about 26% of the time in 6-menus with only one CS, which was less often than optimal (39%). The IS on the other hand was chosen slightly more often than optimal (18% vs. 17%). Finally, the lower priced of the larger CS (the one with three members) was chosen more often than the lower priced of the smaller CS in 6-menus with two CS, (40% vs. 33%), but less often than optimal (62%), and the IS was chosen more often than optimal.

In the aggregate, consumers do not follow a Naive rule since they take account of the presence of a CS by discarding higher priced CS offers. The LPCS was chosen more often than any other offers but there is no consistent evidence across menus length that this was due to consumers avoiding IS offers. Rather, this is consistent with consumers following the SF rule (transferring their initial choice of a HPCS to the LPCS), but not with following the DE rule (whereby the LPCS and the IS would be chosen with about equal frequency).

[10] Looking at menus where consumers performed particularly badly, one finds that they mistakenly chose smaller size options, triangles, options to the end of the lexicographic order, or the LPCS when the IS was actually better. This underlines an important point about the CS rule: while following it maximizes average payoffs for a consumer that is prone to making mistakes, it does *not* lead to the correct choice for each individual choice instance.

[11] Differences across hard and easy menus are not significant and are therefore not reported.

Table 3: Choice frequencies by menu length and presence of a CS

		LPCS	HPCS	IS	LPSCS	HPSCS	MPLCS
3-menu	No CS			33.33%			
	One CS	56.71%	5.86%	37.40%			
6-menu	No CS			16.67%			
	One CS	25.55%	3.52%	17.73%			
	Two CS	40.41%	2.35%	17.33%	32.67%	2.97%	4.27%

Notes: In the case of 6-menus with two CS, the LPCS is the Lower Priced of the Larger CS (the one with three members), the HPCS is the Higher Priced of the Larger CS, and the MPLCS is the Middle Priced of the Larger CS. The LPSCS is the Lower Priced of the Smaller CS (the one with two members) and the HPSCS is the Higher Priced of the Smaller CS. In 6-menus with one CS, the IS choice frequency is calculated by averaging across the four IS offers.

Let us now consider whether consumers benefited from the presence of a CS by looking at individual payoffs by menu length, difficulty and presence of a CS (Table 4).

Table 4: Payoffs by menu length, difficulty and presence of a CS

		Hard menus			Easy menus		
		Mean	Std Dev	N	Mean	Std Dev	N
3-menu	No CS	10.41	0.92	1818	11.02	4.56	1818
	One CS	10.45	0.96	1818	13.34*	3.96	1818
6-menu	No CS	10.14	0.81	1818	11.97	4.11	1818
	One CS	10.04*	0.98	1818	13.84*	5.48	1818
	Two CS	10.78*(*)	0.87	808	12.78*(*)	4.34	808

* Difference significant *vs.* one row above.

(*) Difference significant *vs.* two rows above.

This table can be read in conjunction with another table that indicates how those payoffs translate in terms of how close they are to the maximum available payoff in each menu. Table 5 thus reports the average of the ratio $(\mathrm{up}^{\max} - \mathrm{up}^{\text{chosen}})/(\mathrm{up}^{\max} - \mathrm{up}^{\min})$ over individuals and menus in each category. We normalize the difference between the worst choice and the consumer's choice as shown because we want to be able to compare performance between easy and hard menus, where the difference between the worst and the best choice within a menu will be smaller on average. We call this the performance ratio. A value of 0 would indicate the consumers always made the worst choice, while a value of 1 would indicate they always made the best choice.

Table 5: Performance ratio by menu length, difficulty and presence of a CS

		Hard menus			Easy menus		
		Mean	Std Dev	N	Mean	Std Dev	N
3-menu	No CS	0.597	0.447	1818	0.607	0.448	1818
	One CS	0.592	0.419	1818	0.794*	0.324	1818
6-menu	No CS	0.683	0.353	1818	0.682	0.321	1818
	One CS	0.545*	0.364	1818	0.735*	0.299	1818
	Two CS	0.735*(*)	0.323	808	0.759(*)	0.365	808

* Difference significant *vs.* one row above.

(*) Difference significant *vs.* two rows above.

Subjects obtained a payoff of 11.44 ECU on average (1 ECU=0.8 €), and their performance ratio was 0.66. No participant obtained payoffs that were significantly less than 10.22, which is what they would have obtained had they chosen at random within our menus, and only 8 obtained payoffs that were not significantly greater than this. Subjects therefore seem to have made considered choices. As could be expected from statistical arguments, individuals obtained higher payoffs with 6-menus and with easy menus.

When choosing from menus with no CS, participants obtained 10.89 ECU (std. dev. 3.21) and their performance ratio was 0.64 (std. dev. 0.40), while when choosing from menus with one CS they obtained 11.91 ECU (std. dev. 3.84) while their performance ratio was 0.67 (std. dev. 0.37). Participants thus generally obtained significantly higher payoffs and performed significantly better when a menu included a CS, *except* in the case of hard 6-menus with one CS, where payoff was lower. The presence of a CS did not therefore consistently improve consumer payoffs when menus were hard, but significantly and consistently increased payoffs when menus were easy.

Panel regressions of payoffs on individual and menu characteristics (not reported) indicate that women obtained higher payoffs and subjects with higher scores in the mathematical and practical consumption problems obtained higher payoffs as well. Payoff increased with the order in which the menu was presented so there was some learning. Motivation, scores in the shape comparison task and time spent choosing an offer within each menus did not appear to have a significant effect.[12] There was no individual effect, that is, no individual seemed to perform better than others above and beyond what could be predicted from their gender and scores in control tasks. Easier menu, menus with more choices, and the presence of a CS also increased payoffs. The effects above are robust to various specifications.

When mapping payoffs by menu length and difficulty (rows with no CS in table 4) to the predictions from our simulations (Graph 4), we find that they correspond to a situation in which the standard error of the consumers' error term is 0.15 – though consumers seem more accurate when menus were hard. A tentative explanation may be that consumers could perceive that prices in some menus were closer together than in others, and thus paid more attention in those cases. Lower accuracy when choosing from easy menus did not prevent them from obtaining higher payoffs there

[12] We checked also if there was some quadratic effect in terms of time spent, with time spent increasing payoffs but fastest times (inattention) and slowest times (difficulty) obtaining lower payoffs. While coefficients were of the correct sign, they were not significant.

than when choosing from hard menus however. Consumers did not obtain higher payoffs in hard 6-menus than in hard 3-menus, that is, they appear to have been less accurate when faced with more choice. Note that the optimal threshold v_j if the standard error of the consumer's error term is 0.15 would be between 1.2 and 1.4. We will see that even those consumers who followed the Threshold rule generally chose thresholds that were lower than this, indicating perhaps that they were over-confident in their ability to make accurate choices.

4 Econometric analysis

We first determine in this part how consumers make choices among options in menus with no CS, then consider their choices among menus with one CS, and finally determine rules followed by consumers at the individual level. This will allow us to determine whether indeed consumers prefer offers that are presented in terms of a CS. The menus with no CS are used to simulate the outcome of various choice rules the consumers may follow when faced with menus that include one CS (we do not present the analysis for menus with two CS). Those predictions are then compared with the observed choices to determine what choice rule best predicts consumer choice, at the individual and at the global level. We therefore begin with the expression of the model to predict consumer choice among menus with no CS.

4.1 Consumer choice when there is no common standard

We perform maximum likelihood estimation with three different models, the alternative-specific conditional logit and probit models and the mixed logit model which allows for preference heterogeneity for all the attributes. The probit model is fitted by using maximum simulated likelihood implemented by the Geweke-Hajivassiliou-Keane (GHK) algorithm (Greene and Hensher, 2003). The Halton sequence is used to generate the point sets used in the quasi-Monte Carlo integration of the multivariate normal density, while optimization is performed using the Berndt-Hall-Hall-Hausman procedure (Berndt et al., 1974). The mixed logit model is fitted by using maximum simulated likelihood (Train, 2003) and the estimation was performed with the user-written `mixlogit` command for Stata (Hole, 2007). Estimation makes use of the sandwich estimator of variance, except when performing the probit regressions with 6-menus as convergence was not achieved otherwise.

The outcome for each menu is one of 3 or 6 options. Options are identified by their position in the menu if there is no CS, and by whether they are the LPCS, HPCS or an IS in menus with a CS. The dependent variable is the choice of the consumer among alternatives and the independent variables include the unit price of the option, its shape, its size and its position. Since shapes that extend more broadly in space are preferred (see Krider et al., 2001), we create a variable coding shapes from most to least attractive: a triangle is assigned a value of 1, a square a value of 2 and a circle a value of 3.[13] The variable "position" is coded by lexicographic position in the menu, from 1 if the option is in the top left corner to 6 if it is in the bottom right corner in a 6-menu, otherwise to 3 for the option to the right in a 3-menu. As per a remark in Hole (2007), we include no alternative-specific constants in our models, which is "common practice when the data come from so-called unlabeled choice experiments, where the alternatives have no utility beyond the characteristics attributed to them in

[13] We also ran the same regressions with each shape being a dummy variable. This did not influence the results.

the experiment." We will also cross unit price with case specific variables such as gender and scores in the control tasks to determine whether individual characteristics make our subjects more or less sensitive to price signals (other individual characteristics such as age and educational background do not vary sufficiently in our sample). We also consider a menu specific variable (whether the menu was "hard" or "easy") and variables that are both menu and case specific (the order in which a specific menu was presented to an individual and the time that individual spent deciding on this menu).

Formally, denote y^o_{ijm} the utility of option j in menu m for individual i, and denote $y_{ijm} = 1$ if that option is chosen. We will have $y_{ijm} = 1$ if $y^o_{ijm} > y^o_{itm}$ for all $t \neq j$ in menu m, 0 else. Latent utility y^o_{ijm} takes the form $y^o_{ijm} = \alpha\, up_{jm} + \omega \times up_{jm} \times \Omega_i + \mu \times up_{jm} \times M_m + \beta\,\text{shape}_{jm} + \gamma\,\text{size}_{jm} + \phi\,\text{position}_{jm} + u_{ijm}$ with u_{ijm} a random variable of mean 0 that follows either a logistic or a normal distribution. Ω_i is a $q \times 1$ vector of case-specific variables while ω is a $1 \times q$ vector of parameters. M_m is a $h \times 1$ vector of menu-specific variables while μ is a $1 \times h$ vector of parameters.

We find that a model that takes into account all the alternative specific variables (price, position in menu, shape, area size) minimizes the Akaike Information Criterion ("AIC"). In addition to those, one menu specific variable was consistently significant across menu length (whether the menu was easy or hard) and one case specific variable turned out to be significant for 3-menus (performance in the shape comparison task). Results are shown in table 6. Subjects tend to prefer options that have a lower unit price, "broader" shapes, and smaller sized options (equivalently, those with lower displayed prices). There is no consistent tendency for consumers to favor either options at the beginning or at the end of the menu. Subjects with low performances in the shape comparison task were understandably less affected by unit price in their choice, and subjects were more sensitive to unit price in hard menus.

The log-likelihood is much lower in 6-menus than in 3-menus, which means that the choices from 6-menus are considerably less accurately predicted with our model than from 3-menus (there was the same number of choices to make from within each menu type). This means there is more randomness in consumer choice within 6-menus, probably because it is more difficult to compare 6 offers than 3 offers as this requires holding more information into one's working memory.

Results from the mixed logit model indicate there is significant variation in the extent to which an option's shape and size influenced consumers. However, the influence of an option's position did not appear to vary across subjects. We can conclude that our participants have some bias that may be explained by their use of a budget rule (choose lower priced, that is, smaller sized, options) and of a shape rule (prefer triangles to square to circle). However, the marginal effect of an increase in unit price is much higher than that of any other variables (not reported).

4.2 Consumer choice when there is a common standard

The analysis of the case where there is a CS differs from the case where there is no CS in that options in a menu differ in nature depending on whether they are the LPCS, the HPCS or an IS. Whether a subject avoids the HPCS or prefers the LPCS vs. the ISs may depend on their individual characteristics so that we introduce case-specific variables (here, a case is an individual) along with alternative-specific variables to determine choice among alternatives. Our case specific variables are scores in the mathematical, shape comparison and simple problems, along with gender, time spent choosing within a menu and motivation. We also consider whether facing a hard menu makes it more likely to favor

Table 6: Regressions with no CS, 3 and 6-menus

	(1) Logit 3-menus	(2) Probit 3-menus	(3) MixLogit 3-menus	(4) Logit 6-menus	(5) Probit 6-menus	(6) MixLogit 6-menus
main						
unit price (up)	−18.7200***	−16.2710***	−20.2468***	−16.2815***	−6.6052***	−17.4617***
	(−6.89)	(−6.87)	(−6.91)	(−8.49)	(−6.19)	(−7.67)
up × hard menu	−9.7361**	−11.7958***	−10.2537**	−24.3972***	−9.6439***	−26.9113***
	(−2.86)	(−3.50)	(−3.14)	(−6.63)	(−4.77)	(−7.19)
up × score shape task	20.7627*	14.2704+	20.9114*	10.8997	4.1976*	13.4579
	(2.24)	(1.77)	(2.23)	(1.53)	(1.96)	(1.61)
position	0.0656*	−0.0916+	0.0671**	0.0053	0.0240	0.0046
	(2.56)	(−1.95)	(2.63)	(0.53)	(1.22)	(0.44)
shape	−0.3621***	−0.3705***	−0.3961***	−0.3339***	−0.1509***	−0.3958***
	(−12.05)	(−11.55)	(−9.21)	(−14.52)	(−6.58)	(−9.54)
size	−0.0121***	−0.0108***	−0.0137***	−0.0019	−0.0002	−0.0019
	(−5.28)	(−4.19)	(−4.03)	(−0.92)	(−0.23)	(−0.41)
SD						
shape			0.3836***			0.4549***
			(9.81)			(9.48)
size			0.0352***			0.0537***
			(8.39)			(11.81)
N	10854	10854	10854	21708	21708	21708
ll	−3757.4104	−3747.6265	−3689.0559	−6103.5141	−6042.2092	−5881.8136

t statistics in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: One subject did not perform the shape comparison task, so the regressions are based on 201 subjects choosing among 18 menus with no CS.

the LPCS as following a simple heuristic may be more likely if there appears to be little difference in prices between options. Finally, we consider whether the LPCS was next to the HPCS on the same row in the menu since it is easier to notice there is a CS if CS options are close together.

The model above is thus modified as follows: Latent utility y^o_{ijm} takes the form $y^o_{ijm} = \alpha\, up_{jm} + \omega \times up_{jm} \times \Omega_i + \mu \times up_{jm} \times M_m + \lambda_j \times \Omega_i + \theta_j \times M_m + \beta\,\text{shape}_{jm} + \gamma\,\text{size}_{jm} + \phi\,\text{position}_{jm} + u_{ijm}$. As before, j is the option, m is the menu and i is the individual. An option is coded in terms of whether it is the LPCS, the HPCS or an IS offer. Ω_i is a $q \times 1$ vector of case-specific variables, the same variables being assumed to influence the choice for each option, ω is a $1 \times q$ vector of parameters, M_m is a $h \times 1$ vector of menu-specific variables while μ is a $1 \times h$ vector of parameters. λ_j is a $1 \times q$ vector of parameters, different for each alternative as case-specific variables are assumed not to influence the choice of each alternative in the same way. Similarly, θ_j is a $1 \times h$ vector of parameter translating the influence of menu characteristics on the choice of an alternative. u_{ijm} is a random variable of mean 0 that follows either a logistic or a normal distribution. We constrain λ_j and θ_j to be the same for all four IS options in 6-menus. Model selection using the AIC finds that all of the alternative specific variables ought to be used, while only score in the shape comparison and in the mathematical tasks, along with gender and whether a menu is hard or easy, ought to be used as case-specific variables. Results are reported in table 7.

Table 7: Regressions with a CS, 3 and 6-menus

		(1) Logit 3-menus	(2) Probit 3-menus	(3) MixLogit 3-menus	(4) Logit 6-menus	(5) Probit 6-menus	(6) MixLogit 6-menus	(7) Logit 6-m 2 CS	(8) Probit 6-m 2 CS	(9) MixLogit 6-m 2 CS
main	unit price (up)	−14.5806 * **	−9.5842 * *	−15.5821 * **	−19.2308 * **	−6.9231 * **	−19.6781 * **	−11.6973 * **	−1.9796	−12.2147 * **
		(−4.99)	(−2.71)	(−4.65)	(−11.03)	(−4.85)	(−9.62)	(−4.09)	(−0.28)	(−3.63)
	up × shape task	−5.0589	−4.1344	−5.2897	17.5590 * *	5.8729*	18.0310*	6.4474	1.0936	7.1735
		(−0.48)	(−0.61)	(−0.45)	(2.85)	(2.53)	(2.49)	(0.64)	(0.27)	(0.59)
	up × hard menu	16.4312 * **	10.9950 * *	16.8252 * **	15.0242 * **	4.7505 * **	16.0361 * **	8.8646	−1.2748	10.2190
		(3.90)	(3.29)	(3.44)	(4.72)	(3.44)	(5.17)	(0.88)	(−0.25)	(0.99)
	position	−0.0543+	−0.0463+	−0.0489	0.0696 * **	0.0325 * **	0.0656 * **	0.0205	0.0028	0.0303
		(−1.75)	(−1.94)	(−1.59)	(6.42)	(3.94)	(6.08)	(0.84)	(0.21)	(1.16)
	shape	−0.2062 * **	−0.1465 * *	−0.2333 * **	−0.3682 * **	−0.1332 * **	−0.4286 * **	−0.6777 * **	−0.0766	−0.7630 * **
		(−5.21)	(−2.61)	(−4.96)	(−14.56)	(−4.84)	(−10.18)	(−9.49)	(−0.28)	(−10.17)
	size	−0.0007	−0.0004	0.0009	−0.0058 * *	−0.0019*	−0.0062+	−0.0447 * **	−0.0036	−0.0370 * **
		(−0.21)	(−0.16)	(0.20)	(−3.21)	(−2.04)	(−1.70)	(−5.87)	(−0.28)	(−3.73)
HPCS	score shape task	1.9492*	1.4973 * *	1.9687*	2.7203*	1.1523 * *	2.7527*	−0.1989	0.0525	−0.2176
		(2.44)	(2.73)	(2.12)	(2.48)	(2.78)	(2.45)	(−0.12)	(0.04)	(−0.12)
	score math	−0.0096	−0.0113	−0.0090	0.0675+	0.0239	0.0689+	−0.0776+	−0.0508	−0.0756
		(−0.41)	(−0.69)	(−0.30)	(1.87)	(1.61)	(1.85)	(−1.71)	(−0.97)	(−1.41)
	gender	−0.4832 * *	−0.2882 * *	−0.4819*	−0.8498 * **	−0.3015 * **	−0.8506 * **	−1.3970 * **	−0.8411*	−1.3944 * **
		(−3.24)	(−2.74)	(−2.43)	(−4.30)	(−3.43)	(−3.66)	(−4.06)	(−2.45)	(−3.71)
	hard menu	−0.1681	−0.1401	−0.1995	−0.5105*	−0.0970	−0.5341 * *	−0.3825	0.2909	−0.4529
		(−1.03)	(−0.90)	(−1.34)	(−2.55)	(−1.08)	(−2.58)	(−0.94)	(0.61)	(−1.10)
	close CS	−0.4200+	−0.1512	−0.4122*	−0.1289	0.0471	−0.1133	−0.2745	−0.4255	−0.2341
		(−1.95)	(−0.97)	(−1.98)	(−0.50)	(0.46)	(−0.45)	(−0.78)	(−1.03)	(−0.61)
	constant	−1.5065*	−1.7151 * **	−1.4918*	−2.9550 * *	−1.7752 * **	−2.9806 * *	0.2745	−1.1791	0.2397
		(−2.42)	(−3.67)	(−1.97)	(−3.25)	(−3.40)	(−3.23)	(0.22)	(−0.73)	(0.16)
IS	score shape task	−0.2998	−0.3094	−0.4166	0.0294	0.1489	0.0205	0.1574	0.0123	0.1212
		(−0.77)	(−1.07)	(−0.94)	(0.06)	(0.87)	(0.04)	(0.18)	(0.11)	(0.15)
	score math	0.0248+	0.0188+	0.0284	0.0054	0.0016	0.0062	0.0078	0.0004	0.0103
		(1.84)	(1.74)	(1.54)	(0.37)	(0.27)	(0.36)	(0.29)	(0.13)	(0.30)
	gender	−0.2679 * **	−0.1628*	−0.3129 * *	−0.2517 * *	−0.1132 * *	−0.2657 * *	−0.3562*	−0.0368	−0.3590*
		(−3.45)	(−2.23)	(−3.18)	(−2.80)	(−2.78)	(−2.69)	(−2.30)	(−0.28)	(−2.25)
	hard menu	0.0105	−0.0190	−0.0386	−0.4933 * **	−0.1580 * **	−0.4923 * **	0.6969 * *	0.0083	0.5104*
		(0.13)	(−0.34)	(−0.45)	(−5.46)	(−3.43)	(−5.51)	(3.03)	(0.15)	(1.96)
	close CS	−0.7851 * **	−0.5384 * *	−0.8847 * **	0.2557*	0.0519	0.2872 * *	0.4785 * *	−0.0070	0.6749 * *
		(−6.39)	(−2.94)	(−7.38)	(2.36)	(1.16)	(2.65)	(2.65)	(−0.22)	(3.21)
	close SCS							−0.2762	0.0985	−0.2752
								(−0.69)	(0.27)	(−0.71)
	constant	0.0291	−0.0573	0.0736	0.1316	0.0202	0.0741	−1.2370+	−0.0573	−1.3206
		(0.08)	(−0.23)	(0.15)	(0.35)	(0.13)	(0.17)	(−1.73)	(−0.25)	(−1.53)
SD	shape			0.3763 * **			0.4722 * **			0.2567*
				(7.63)			(10.60)			(2.38)
	size			0.0363 * **			−0.0428 * **			0.0650 * **
				(8.18)			(−11.42)			(7.39)
	N	10851	10851	10851	21708	21708	21708	9648	9648	9648
	ll	−2919.5938	−2917.2251	−2850.6984	−5617.2643	−5564.1585	−5450.2711	−2078.6207	−2057.6991	−2055.1065

t statistics in parentheses + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Base outcome is the LPCS. Not reported in the table are coefficients on case specific variables for the higher and lower priced of the options that are part of the CS with two options and for the middle priced of the CS with three options in 6-menus with two CS. Whether CS options were close within a CS was deemed to influence choice only within that CS and vs. the IS option.

In terms of alternative-specific variables, our results when there was no CS are confirmed, that is, subjects tend to prefer lower priced options, "broader" shapes, and smaller sized options (equivalently, those with lower displayed prices). One can notice that prices being close together (hard menus) makes consumers less sensitive to price, the opposite to the result when there is no CS. This might be due to consumers relying more on the CS rule when there is little apparent difference between offers: this rule can lead them to choosing the LPCS even when its price is actually higher than that of the IS. This hypothesis receives some support from how consumers are more likely to avoid the IS in 6-menus that are hard.

Case-specific variables show that consumers tend to avoid the HPCS: the parameter on the constant term for that option is negative and highly significant. Individuals that are worst at the shape comparison tasks are more likely to choose the HPCS, maybe because they find it difficult to compare the area and shape of all options and thus do not notice the presence of a CS. It is however only women who display an aversion to the IS *vs.* the LPCS. This aversion to the IS is encouraged when the presence of a CS is more obvious, i.e. when the CS options are next to each other – there was a negative impact of the dummy variable taking value one if CS options are close in 3-menus (the impact was not consistent across logit and probit regressions in the case of 6-menus). Whether the menu is hard also encourages individuals in rejecting the IS option, at least in 6-menus with one CS (results in the case of 6-menus with two CS are not consistent across logit and probit regressions).

In conclusion, only women appear to follow the CS rule when choosing among options. This might explain why women managed to obtain higher payoffs than men in this experiment even though they were less good at control tasks that predict higher payoffs.

5 Assignment of consumers to rules

We seek in this part to compare consumers' decisions with what would be predicted under different selection rules, as presented in section 2.[14] Consider table 8 which shows the unit price and standard of three options in a menu, the choice of a consumer among those three options, and the predicted choice under different choice rules. Predicted choice is expressed in terms of probabilities. The consumer is said to follow the choice rule he is closest to in terms of log likelihood, weighted by the number of degrees of freedom allowed by each rule.

[14] We do not consider the alternative selection rules exposed in section 2.4 since the biases they reflect are already corrected for in our regressions via the shape, size and position variables, and the effect of those variables was relatively small.

Table 8: Rule predictions and consumer choice

	HPCS	LPCS	IS
Unit price	0.51	0.50	0.48
Standard	A	A	B
Consumer choice	0	1	0
Naive	p_{HPCS}^{Na}	p_{LPCS}^{Na}	p_{IS}^{Na}
CS	0	1	0
SF	0	$p_{HPCS}^{Na} + p_{LPCS}^{Na}$	p_{IS}^{Na}
DE	0	$p_{LPCS}^{Na}(LPCS, IS)$	$1 - p_{LPCS}^{DE}$
Threshold rule	0	$p_{LPCS}^{Na}(LPCS, IS \times v)$	$1 - p_{LPCS}^{Th}$

In the case shown in table 8, the consumer chooses the LPCS, and may thus be following any of the possible rules. If he had chosen the HPCS, then one would have been able to say he must be Naive. If he had chosen the IS, then one could have said he was not following the CS rule. If the same type of choice is offered several times and the consumer never chooses the HPCS, then he is unlikely to be Naive. If he always choose the LPCS, then he is likely to follow the CS rule. If he chooses either the LPCS or the IS, but more often the LPCS, then he is likely to follow either the SF rule or a Threshold rule.

We use the estimation results from the mixed logit regressions done for the case where there is no CS to predict choice when there is a CS. If the consumer is Naive, his choice will be predicted by applying parameter estimates from the model with no CS to the data with CS. If he follows the CS rule, he will choose the LPCS. If he follows the SF rule, then the probability to choose the LPCS is the sum of the probability to choose the HPCS or the LPCS obtained when applying parameter estimates from the model with no CS. If he follows the DE rule, then the HPCS is excluded from the consideration set (this is done by applying the parameter estimates from the model with no CS to the data with a CS modified such that the price of the HPCS is unaffordable). An issue is that the CS rule predicts the IS will never be chosen, which means that any consumer who ever chose an IS even if he always chose the CS otherwise would be predicted not to follow the CS by the maximum likelihood criterion. We can confirm that no consumer systematically chose the LPCS within every menus with a CS. Therefore, strictly speaking, no consumer followed the CS rule. This is where the *Threshold rule*, which spans the gap between the CS and the DE rule, comes into play.

In terms of notation, we denote the probability the LPCS is chosen under the DE rule as $p_{LPCS}^{DE} = p_{LPCS}^{Na}(LPCS, IS)$. This is to be interpreted as the probability a Naive consumer would choose the LPCS if his choice was restricted to either the LPCS or the IS. Similarly, the probability the LPCS is chosen under the Threshold rule is $p_{LPCS}^{Th} = p_{LPCS}^{Na}(LPCS, IS \cdot v)$. This is to be interpreted as the probability a Naive consumer would choose the LPCS if his choice was restricted to either the LPCS or the IS and the price of the IS was multiplied by a factor v. We computed for each consumer the threshold v_j that maximizes their maximum likelihood. Subjects with a high value of v_j are close to following the CS rule, while those with low v_j are close to following the DE rule.

The CS, DE and SF rules predict that the HPCS will never be chosen. However, as we saw, this is not the case in our data. One therefore has to take account that some consumers choose the HPCS. We therefore do a separate regression so as to determine the probability p_{LPCS} with which the LPCS is chosen among CS offers. Note that in this case, only the offer's position and its price may determine the choice, along with some case-specific variables, since both shape and area are the same in a CS. One then modifies the formulas above as follows: In the case of the Signal-First rule: $p^{SF}_{LPCS} = p_{LPCS}(p^{Na}_{HPCS} + p^{Na}_{LPCS})$ while $p^{SF}_{HPCS} = (1 - p_{LPCS})(p^{Na}_{HPCS} + p^{Na}_{LPCS})$ and p^{SF}_{IS} is as before. In the case of the CS rule: $p^{CS}_{LPCS} = p_{LPCS}$ and $p^{CS}_{HPCS} = 1 - p_{LPCS}$. In the case of the DE rule: $p^{DE}_{LPCS} = p_{LPCS} p^{Na}_{LPCS}(LPCS, IS)$, $p^{DE}_{HPCS} = (1 - p_{LPCS}) p^{Na}_{HPCS}(HPCS, IS)$ and $p^{DE}_{IS} = 1 - p^{DE}_{LPCS} - p^{DE}_{HPCS}$. The principle is the same for obtaining p^{Th}. Formulas are slightly longer in the case of 6-menus but can be inferred from the above.

Compared to the Na predictions, both the SF and the DE predictions make use of an additional degree of freedom as they require CS information. Compared to the SF and the DE, the Threshold rule makes use of yet one more degree of freedom as it requires estimating the threshold used by the subjects. This will be taken into account by comparing rules using the Akaike Information Criterion ("AIC").

In mathematical terms, the likelihood function is $f(y,\theta) = \prod_{t=1}^{N} \prod_{j=1}^{M} p_{tj}^{y_{tj}}$ with t denoting the menu, N the total number of menus presented to consumers, j denoting the option, M the number of options, and $y_{tj} = 1$ iff $y_t = j$, 0 otherwise, whereby y_t is the consumer's choice. $p_{tj} = \Pr(y_t = j)$ is the predicted probability, which depends on the rule we assume for consumers' choice, so for example $p_{tj} = 1$ iff j is the LPCS and the consumer is assumed to follow the CS rule. y is the vector of choices and θ are the parameters determining the choice among options.

5.1 What rule best describes aggregate behavior?

Table 9 reports the log-likelihood, the values of the AIC and of the Bayesian information criterion ("BIC") for each rule, for 3 and 6-menus.[15] The last column contains the value of threshold v that maximizes the log-likelihood for the Threshold rule. The number v reported there is to be interpreted as "consumers appear to consider IS options as v times more expensive when they are presented next to CS options than when they are presented next to other IS options". This measures the price penalty applied to IS options when compared to the LPCS. For more interpretation of this number, see the detailed explanation in section 2.2.

[15] We only study 6 menus with one CS.

Table 9: Rules scores, aggregate behavior

		Naive	Signal First	Dominance Editing	Threshold Heuristic	v
3-menus	LL	−3484	−2994	−3073	−2984	1.05
	df	8	9	9	10	
	AIC	6984	6006	6164	5988	
	BIC	7034	6062	6220	6050	
	N	3618	3618	3618	3618	
6-menus	LL	−5954	−5769	−5788	−5762	1.04
	df	8	9	9	10	
	AIC	11924	11556	11594	11544	
	BIC	11974	11612	11650	11606	
	N	3618	3618	3618	3618	

The threshold heuristic gives the best predictions for both menu lengths, while the Signal-First heuristics comes second according to the AIC. The Naive rule is clearly rejected in all cases so consumers do take CS information into account. In terms of threshold, an IS offer suffers a 4 to 5% price penalty compared to the LPCS offer, which is a considerable amount. The consumers do not in the aggregate appear to merely follow the dominance editing heuristic, that is, they do tend to disfavor IS offers in favor of the LPCS.

While those aggregate results are interesting in their own right, we are more interested in individual behavior so we attempt to determine rules followed by individuals in the next section.

5.2 What rules do individuals follow?

The above techniques were used to determine rules followed by the subjects. Table 10 cross-tabulates the number of subjects assigned to each type when looking at 3-menus and at 6-menus:

Table 10: Subjects assigned to rules, by menu length.

		6 menus				
		Naive	SF	DE	Th	Total
3-menus	Naive	27	5	5	1	38
	SF	26	19	23	3	71
	DE	19	15	23	4	61
	Th	9	10	9	3	31
	Total	81	49	60	11	201

Pearson's chi-square test of independence rejects the hypothesis that types are independent between 6- and 3-menus. Subjects thus tend to follow the same rules in both menu-lengths, that is, the numbers in the diagonal of the table tend to be the highest of their respective rows and columns. When not keeping to the same rules, those who followed the DE or the SF rule in 3-menus tend to become Naive when choosing among 6-menus. The higher number of subjects being Naive when faced with 6-menus (81 *vs.* 38 in 3-menus) tends to confirm that subjects do not notice the presence of a

CS in 6-menus. Worse, the lower number of subjects following the SF rule in 6-menus *vs.* 3-menus (49 *vs.* 71 in 3-menus) means that subjects may not even realize, after making a provisional choice that is a CS, that it *is* a CS.

In terms of payoffs, and whether considering 3-menus or 6-menus, consumers following the Naive rule tend to obtain significantly lower payoffs than consumers of all other types, as could be expected from our analysis of rule performance in section 2.3. Consumers that follow the DE and the SF rules obtain comparable payoffs and obtain significantly higher payoffs than Naive consumers. Those who follow the Threshold rule also obtain significantly higher payoffs than Naive consumers but do not obtain significantly higher payoffs than either the DE or the SF.

However, this is because some consumers, rather than favoring CS, actually disfavor them, and are thus assigned to the Threshold rule as well. When considering only the 27 subjects of the 31 assigned to the Threshold rule that do favor CS in 3-menus, and the 6 subjects of the 11 assigned to the Threshold rule that do favor CS in 6-menus, their payoffs are not significantly higher than the payoffs of DE and SF subjects when considering 3-menus, but significantly higher when considering 6-menus.

Overall therefore, consumers that do not realize there is a CS do tend to obtain lower payoffs than others, those who follow the SF, DE and Threshold rule obtain comparable payoffs in 3-menus, while those following the Threshold rule perform significantly better in 6-menus. While theory would have predicted that consumers following the SF rule would obtain higher payoffs than those who follow the DE rule, the difference between their payoffs was not predicted to be large, explaining perhaps the lack of significance of the difference in payoffs between the two.

5.3 Do consumers choose their threshold rationally?

In terms of thresholds used by those individuals that were assigned to the Threshold rule, theory presented in this paper would predict that a rational consumer who is beset by an inability to assess offers accurately ought to be using higher thresholds than those used by subjects that are more accurate. Accuracy can be estimated by the payoffs consumers obtained when faced with menus with no CS. Those who obtained higher average payoffs in those menus are more accurate. The following graphs relate average payoffs obtained by subjects in menus with no CS to the threshold they used when choosing within menus with a CS. Bigger points indicate those individuals that were actually determined to be following the Threshold rule. We super-impose on this graph the optimal choice of threshold for a consumer with accuracy implied by his average payoff when faced with menus with no CS.[16] The graph for 3-menus can be read as follow: Consider point (105,11), which represents a consumer who obtained a payoff of 11 in 3 menus with no CS and used a threshold of 1.05 in 3 menus with a CS. The curve indicates that a threshold of 1.20 would have been the optimal choice for this consumer.

[16] We computed the optimal threshold to be used when the consumer knows the distribution of price variances across menus but does not know, when presented with a menu, whether the menu has high or low price variance, as this seems more reasonable to us. That is, with reference to part 2.2, expectation in the formula determining v_j^* is taken over all menus of a specific length.

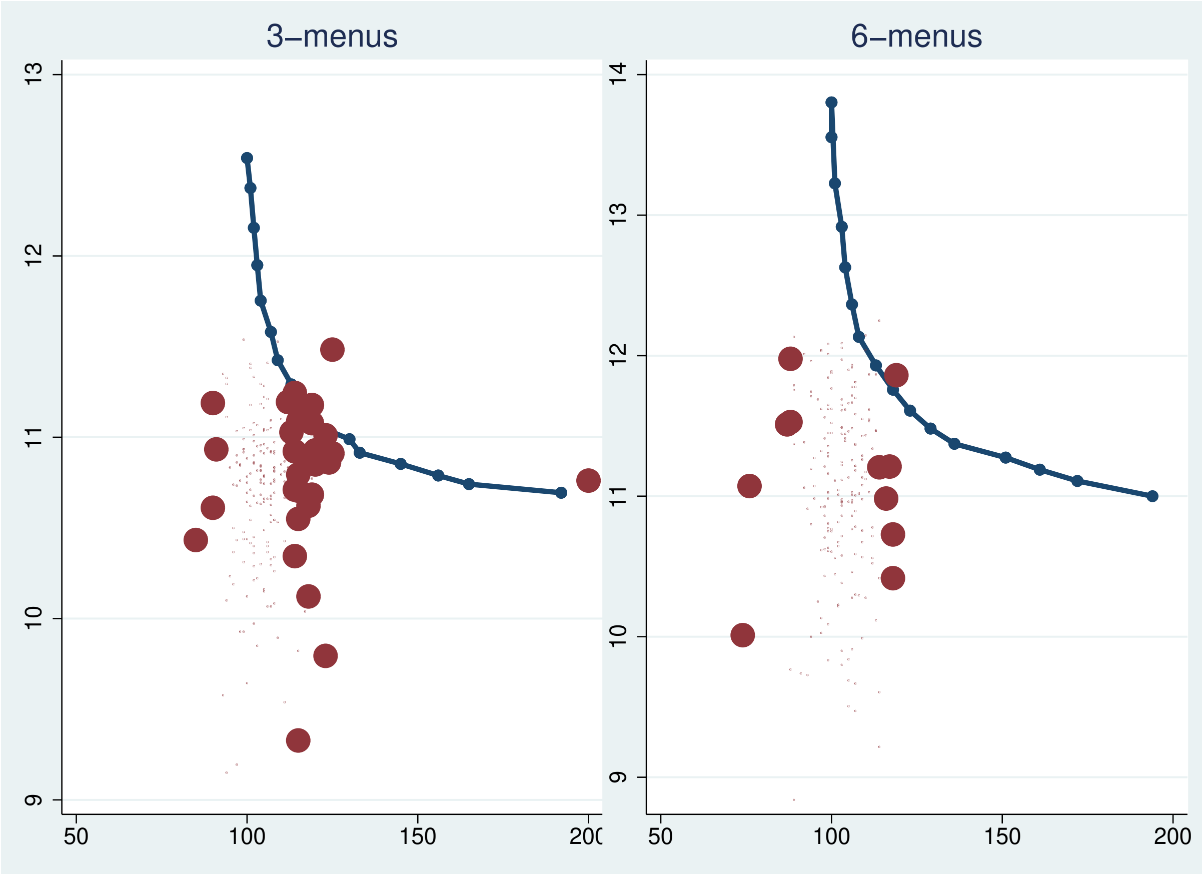

Figure 5: Optimal vs. realized threshold in 3 and 6-menus

We find no relation between payoffs when faced with menus with no CS and the threshold used by the consumer, and this whether the consumer was predicted to be following the Threshold rule or not. Furthermore, almost all points on the graph above are below the optimal line, meaning that consumers adopted lower thresholds than optimal. Consumers who favored the CS adopted thresholds of average 1.17 in 6-menus and 1.20 in 3-menus, while they would have benefited from adopting much higher thresholds generally. While consumers that use the Threshold rule can be considered as being more *savvy* than those who follow other rules, they are therefore not fully rational: they either are not aware of their own level of accuracy, or do not make the link between their accuracy and the threshold they ought to be using. Alternatively, consumers may have too much confidence in their own ability to choose the best offers so they tend to use thresholds that are too low.[17]

5.4 Do consumers learn to follow the Threshold rule?

In this section, we want to see if consumers learn to follow the Threshold rule over the course of the experiment. We therefore exclude the first 20 menus each consumer was faced with – since menus were presented to each consumer in a random order, this will not be the same set of menus for each subject – and run the same regressions as in the above parts. Assignments to types are reported in table 11 below:

[17] In that respect, we note that following the DE rule can be seen as consistent if the consumer is fully confident in his ability to assess prices accurately.

Table 11: Subjects assigned to rules, by menu length, minus first 20 menus

		6 menus				
		Naive	SF	DE	Th	Total
3-menus	Naive	17	8	15	3	43
	SF	26	19	16	9	70
	DE	13	15	22	2	52
	Th	8	12	11	5	36
	Total	64	54	64	19	201

Pearson's chi-square test of independence cannot reject the hypothesis that types are independent between 6- and 3-menus in this case. However, classification is consistent *within* menu lengths, that is, 65% of individuals are classified the same way in 3-menus, and 78% in 6-menus, whether one considers all menus or one excludes the first 20 menus. This indicates that while our classification by type is robust within menus of the same length, this is not so across menu-length, indicating perhaps that individuals do not follow the same rules in both cases. In terms of learning, more individuals used the Threshold rule in later stages of the experiment, especially when considering 6-menus, but the increase was small: only 31 disfavored the IS in 3-menus and 10 in 6-menus, *vs.* 27 and 6 respectively when considering the whole sample. Learning in such an environment, where no menu was ever the same, would in any case be rather difficult, since it is difficult for subjects to ascribe a high payoff (subjects were told their payoff after each try) to the strategy they followed or to the menu having offered a good opportunity. However, the number of subjects following a Naive selection rule in 6-menus decreased when excluding the first 20 tries, which means subjects learned to be more careful in looking out for a CS in 6-menus.

6 Conclusion

We found that as many as 15% of consumers favored CS offers when choosing among 3-menus, but less than 5% did so when choosing among 6-menus, even after gaining experience with this task. *Savvy* consumers (those who favored the CS) applied a large penalty to offers that were expressed in terms of an individuated standard. For those consumers, an offer expressed in terms of an individuated standard had to be 17 to 20% less expensive than an offer expressed in terms of a common standard before they would be as likely to choose it than if it was the best deal among CS offers. In other terms, the best deal among common standard could be priced at \$1.17 and still be *as likely to be chosen* by one of our *savvy* consumers than if it was simply an IS offer priced at \$1.00. Not only is this the case, but this is a perfectly rational behavior when the consumer is not able to determine the value of a product with sufficient certainty. It is then best to be very sceptical of those offers that appear advantageous but are difficult to assess in terms that are comparable with others. However, even those consumers who followed the Threshold rule were not sceptical enough about IS offers: they adopted thresholds that were lower than would have been optimal given their measured inability to make accurate choices among offers.

Even though relatively few consumers followed the Threshold rule, and those who did so adopted lower thresholds than optimal, consumers' aggregate behavior favored offers that were expressed in terms of a common standard. The large penalty applied by savvy consumers to IS offers was only partially diluted into an overall *penalty of 4 to 5% for IS offers* when considering all consumers. This is so even though many consumers followed other rules because the Naive rule is neutral with respect to the existence of a common standard, the Signal-First rule merely re-allocates choice from the higher priced CS offer to the lower priced, and the Dominance Editing rule does not penalize CS offers.

We saw that fewer consumers applied a penalty to IS offers in 6-menus than in 3-menus, and some even applied a penalty to CS offers. This translated in the best CS offer being chosen less often than optimal in 6-menus (see discussion of table 3). We therefore believe that our results are less robust when applied to 6-menus than to 3-menus. It might be that "too much choice" works towards negating the common standard effect, either because it makes it more difficult for consumers to identify offers that are expressed in terms of a common standard, or because they are less confident in the logic of favoring just one offer, the lower priced of the CS offers, when that means disregarding many IS options. This means that while the Threshold rule may be effective in fighting against the introduction of *spurious complexity* by those firms that wish to confuse consumers, it may not be effective in counteracting the introduction of *spurious variety*, whereby firms would pursue what we could call *frame proliferation* when faced with the threat of the emergence of a common standard. For a common standard effect to be effective in markets where there is a multiplicity of choices, firms ought therefore to be able to advertise their use of a common standard. This is where complications occur since the claim to be following a "common standard" may be difficult to verify and there are myriads of ways in which a standard can be debased. For example, if the common standard is in terms of the dimension of the product's packaging, then firms might decide not to fill it properly. If it is in terms of weight, and in the case of food, then managers may lower the quality of the product and mask this by adding more spices. There is therefore a role for regulatory authorities that promote and monitor the use of standards and mandate the disclosure of the information that enters into the definition of that standard.

We would like to examine in future work whether the extent of consumers' preference for CS that was uncovered in this paper is enough to drive a process of convergence towards the adoption of common standards by firms in a competitive market. Our data is encouraging in that respect (Table 12). The lower priced of the CS offers in our menus generated significantly higher revenue than offers expressed in terms of an individuated standard because lower priced CS offers were chosen more often by consumers. Indeed, revenue from a LPCS offer was 0.27 on average when there were three options, much more than revenue of 0.18 for IS offers, and 0.12 on average when there were six options and one CS, much more again than revenue of 0.09 for IS offers. Those differences were significant in a statistical sense as well.

Table 12: Revenue by menu length and presence of a CS

		LPCS	HPCS	IS	LPSCS	HPSCS	MPLCS
3-menu	No CS			0.1643			
	One CS	0.2725	0.0298	0.1786			
6-menu	No CS			0.0816			
	One CS	0.1190	0.0176	0.0860			
	Two CS	0.1891	0.0121	0.0875	0.1570	0.0154	0.0212

Note: See table 3 for the meaning of the headers in the case of 6-menus with two CS.

Since LPCS offers generated significantly higher revenues than others, a firm would prefer to adopt a CS in our setting. Furthermore, consumers who favored CS offers gained higher payoffs than others. While this may not necessarily translate into a process of convergence to a CS as hypothesized in Gaudeul and Sugden (2011), the conditions are in place for this to be so.

References

Adams, S. (1997). *The Dilbert Future.* HarperCollins. (ref. p. 2).

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, and D. D. Evanoff (2010, May). Learning to cope: Voluntary financial education and loan performance during a housing crisis. *American Economic Review 100*(2), 495–500. (ref. p. 2).

Agarwal, S. and B. Mazumder (2010). Cognitive abilities and household financial decision making. Working Paper 2010-16, Federal Reserve Bank of Chicago. (ref. p. 2).

Ariely, D. (2008). *Predictably Irrational.* HarperCollins. (ref. p. 2 and 9).

Ayal, A. (2011). Harmful freedom of choice: Lessons from the cellphone market. *Law and Contemporary Problems 74*, 91–133. (ref. p. 2).

Bar-Gill, O. and R. Stone (2009). Mobile misperceptions. *Harvard Journal of Law & Technology 23*(1), 49–118. (ref. p. 2).

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement 3*(4), 653–655. (ref. p. 17).

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review 151*(3), pp. 1211–1254. (ref. p. 2).

Carlin, B. I. (2009, March). Strategic price complexity in retail financial markets. *Journal of Financial Economics 91*(3), 278–287. (ref. p. 2).

Chioveanu, I. and J. Zhou (2009). Price competition and consumer confusion. *MPRA Paper No. 17340.* (ref. p. 2).

Choi, J. J., D. Laibson, and B. C. Madrian (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies 23*(4), 1405–1432. (ref. p. 2).

DellaVigna, S. and U. Malmendier (2006, June). Paying not to go to the gym. *American Economic Review 96*(3), 694–719. (ref. p. 2).

Eaton, J. W. (2002). *GNU Octave Manual.* Network Theory Limited. (ref. p. 1 and 11).

Ellison, G. (2005, May). A model of add-on pricing. *The Quarterly Journal of Economics 120*(2), 585–637. (ref. p. 2).

Ellison, G. (2006). Bounded rationality in industrial organization. In T. P. Richard Blundell, Whitney K. Newey (Ed.), *Advances in economics and econometrics: theory and applications. Ninth World Congress of the Econometric Society*, Volume 2. Cambridge University Press. (ref. p. 2).

Ellison, G. and S. F. Ellison (2009). Search, obfuscation, and price elasticities on the Internet. *Econometrica 77*(2), 427–452. (ref. p. 2).

Freedman, D. H. (2010). *Wrong: Why experts keep failing us.* Little, Brown and Company. (ref. p. 2).

Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *The American Economic Review 88*(4), 933–946. (ref. p. 9).

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games.* MIT Press. (ref. p. 9).

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540. (ref. p. 2).

Garrod, L., M. Hviid, G. Loomes, and C. W. Price (2008). Assessing the effectiveness of potential remedies in consumer markets. Report 998, Office of Fair Trading. (ref. p. 2).

Gaudeul, A. and R. Sugden (2007). Spurious complexity and common standards in markets for consumer goods. Working Paper 07-20, Centre for Competition Policy, University of East Anglia. (ref. p. 9).

Gaudeul, A. and R. Sugden (2011). Spurious complexity and common standards in markets for consumer goods. *Economica.* Forthcoming. (ref. p. 2, 10, and 30).

Gigerenzer, G. and H. Brighton (2009). Homo Heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science 1*, 107–143. (ref. p. 9).

Gigerenzer, G. and D. G. Goldstein (1999). Betting on one good reason: The take the best heuristic. In G. Gigenrenzer, P. M. Todd, and The ABC Research Group (Eds.), *Simple heuristics that make us smart*, Chapter 4, pp. 75–95. Oxford University Press. (ref. p. 9).

Greene, W. H. and D. A. Hensher (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological 37*(8), 681 – 698. (ref. p. 17).

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal 7*(3), 388–401. (ref. p. 17).

Huber, J. and C. Puto (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research 10*(1), 31–41. (ref. p. 9).

Huck, S. and B. Wallace (2010). The impact of price frames on consumer decision making. Report 1226, Office of Fair Trading. (ref. p. 2).

Iyengar, S. S., G. Huberman, and W. Jiang (2004). How much choice is too much?: Contributions to 401(k) retirement plans. *Pension design and structure New lessons from behavioral finance 401*, 83–95. (ref. p. 2).

Iyengar, S. S. and M. R. Lepper (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology 79*, 995–1006. (ref. p. 2).

Kalaycı, K. and J. Potters (2011). Buyer confusion and market prices. *International Journal of Industrial Organization 29*(1), 14 – 22. Special Issue: Experiments in Industrial Organization. (ref. p. 2).

Krider, R. E., P. Raghubir, and A. Krishna (2001). Pizzas: π or square? Psychophysical biases in area comparisons. *Marketing Science 20*(2), 405–425. (ref. p. 5, 13, and 17).

Lusardi, A. (2008). Financial literacy: An essential tool for informed consumer choice? Working Paper 14084, National Bureau of Economic Research. (ref. p. 2).

Miravete, E. J. (2003). Choosing the wrong calling plan? Ignorance and learning. *American Economic Review 93*(1), 297–310. (ref. p. 2).

Miravete, E. J. (2011). Competition and the use of foggy pricing. *American Economic Journal - Microeconomics*. Forthcoming. (ref. p. 2).

Morwitz, V., E. Greenleaf, and E. J. Johnson (1998). Divide and prosper: Consumers' reaction to partitioned prices. *Journal of Marketing Research 35*, 453–463. (ref. p. 2).

Nunes, J. C. (2000). A cognitive model of people's usage estimations. *Journal of Marketing Research 37*(4), 397–409. (ref. p. 2).

Piccione, M. and R. Spiegler (2010). Price competition under limited comparability. Working paper, LSE, Tel Aviv University and UCL. (ref. p. 2).

Shestakova, N. (2011). Understanding consumers' choice of pricing schemes. Working paper, University of Vienna. (ref. p. 2).

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics 69*(1), 99–118. (ref. p. 13).

Spiegler, R. (2011). *Bounded rationality and industrial organization*. Oxford University Press. (ref. p. 2 and 3).

StataCorp (2009). *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP. (ref. p. 1).

Sugden, R. (1986). *The Economics of Rights, Cooperation and Welfare*. Blackwell. (ref. p. 9).

Sugden, R. (1989). Spontaneous order. *The Journal of Economic Perspectives 3*(4), 85–97. (ref. p. 9).

Sugden, R. (2004). The opportunity criterion: Consumer sovereignty without the assumption of coherent preferences. *American Economic Review 94*(4), 1014–1033. (ref. p. 2).

Thaler, R. H. and C. R. Sunstein (2008). *Nudge*. Yale University Press. (ref. p. 2).

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press. (ref. p. 17).

van Rossum, G. (1995, April). Python reference manual. CWI Report CS-R9525. (ref. p. 1 and 8).

Viswanathan, M., J. A. Rosa, and J. E. Harris (2005). Decision making and coping of functionally illiterate consumers and some implications for marketing management. *Journal of Marketing 69*, 15–31. (ref. p. 2 and 12).

Wilson, C. M. and C. W. Price (2010). Do consumers switch to the best supplier? *Oxford Economic Papers*. (ref. p. 2).

Zeithaml, V. A. (1982). Response to in-store price information environments. *The Journal of Consumer Research 8*(4), 357–369. (ref. p. 2).