JENA ECONOMIC RESEARCH PAPERS

2011 – 044

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

by

Paolo Crosetto
Alexia Gaudeul

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

DO CONSUMERS PREFER OFFERS THAT ARE EASY TO COMPARE? AN EXPERIMENTAL INVESTIGATION*

Paolo Crosetto† and Alexia Gaudeul‡

October 11, 2011

Abstract

Consumers make mistakes when facing complex purchasing decision problems but if at least some consumers choose only among offers that are easy to compare with others then firms will adopt common ways to present their offers and thus make choice easier (Gaudeul and Sugden, 2011). We design an original experiment to identify consumers' choice heuristics in the lab. Subjects are presented with menus of offers and do appear to favour offers that are easy to compare with others in the menu. While not all subjects do so, this is enough to deter firms from introducing spurious complexity in the way they present products.

Keywords: Bounded Rationality, Cognitive Limitations, Common Standards, Consumer Choice, Experimental Economics, Heuristics, Libertarian Paternalism, 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 Modelling 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 and at the 2011 European Conference of the Economic Science Association in Luxembourg in September 2011. The experiment's interface and randomized menu generation were programmed with Python (van Rossum and Drake, 2001). Data analysis and regressions were performed with Stata (StataCorp, 2009). Simulations were run with Octave (Eaton, 2002).

†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

Consumers often go to great length to identify good offers and make thoughtful purchase decisions but still make mistakes when facing complex purchasing decision problems. 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; Lusardi et al., 2010; Miravete, 2003; Wilson and Price, 2010). Knowing this, firms may benefit from introducing spurious complexity in their contract offerings, and thus deliberately obfuscate consumer choice (Carlin, 2009; Chioveanu and Zhou, 2009; DellaVigna and Malmendier, 2006; Ellison, 2005; Ellison and Ellison, 2009; Gabaix and Laibson, 2006; Piccione and Spiegler, 2010; Miravete, 2011).

Adams (1997) refers to sectors in which such firms operate as "confusopolies". He defines this as "a group of companies with similar products who intentionally confuse consumers instead of competing on price". Sectors in which those firms operate 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. This undermines consumers' ability to make informed choices about their services and products.

Gaudeul and Sugden (2011) assert that if at least some consumers limit their choices to those offers that are easily comparable then firms will be forced into simplifying their offers and adopting common standards, thus competing à la Bertrand based on their genuine product characteristics. Offerings that are easily comparable are said to adhere to a common standard ("CS")[1]. The "common standard rule" consists in favoring those. The way this rule fosters competition among firms is called the "common standard effect". This effect is similar to that evoked by Scitovsky (1950) when he mentioned that "it is only the expert buyer who insists on comparing rival products before every purchase; and it is only his insistence on making comparisons that forces the seller - or rather makes it profitable for him - to make his product easily comparable to competing products."

The concept of a "standard" in our setting is essentially the same as what others call 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, a "language" in which a contract is written, but also a unit of measurement, a way of packaging a product, a technical standard, *etc.*..

Note that firms adopting a common standard in the way in which they present their offers do *not* inherently make those offers 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 in its relation with other offers that a CS

[1] An index of abbreviations and notations is provided in appendix A.

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*. Note also that the common standard effect will come into play as soon as some of the aspects of competing products are comparable, that is, it does not require that competing products be comparable across all of their characteristics.

The common standard rule ("CS rule") is a rule of thumb that consumers may use in their selection of which product to buy. However, this rule of thumb is the result of an optimizing process, that is, it is rational as long as there is little difference between products and in the prior beliefs of consumers on the value of each firms' offering. As we will see later, it derives strength from its simplicity, has strong behavioural foundations and can be applied in many settings, thus ensuring its evolutionary robustness. Contributing to the later, there is no need for others to follow it for it to be optimal.

We wish in this paper to determine if some consumers indeed follow such a "common standard rule". Since data from the field would be insufficient for our purpose, we design an experiment to identify what choice heuristics consumers follow when faced with simple choice situations. Subjects are asked to buy paint to cover a fixed area, known and the same for all, and are presented with menus of offerings, each offer being a price and the area the paint can cover for that price. The area is presented in different shapes (circles, triangles and squares) of different sizes. The offers' unit prices are thus difficult to assess correctly (Krider et al., 2001), but offers that are of the same shape and size (of the same "standard") are easy to compare. We find that our subjects differ in their ability to make accurate choices from the menus presented to them but that they generally make better choices (in terms of payoffs) when a menu includes a common standard than when it does not. We observe that many consumers appear to favour the lowest priced of the CS offers ("LPCS"), meaning that they choose it more often than if it had not adhered to the CS. While not all consumers follow the CS heuristic, their number is sufficient to ensure that a firm that would make LPCS offers would make higher profits than firms not adhering to the common standard, meaning that firms have an incentive to adopt a common standard.

Outline: The paper is divided into six sections. We first introduce its context by evoking the debate over what to do about firms wilfully confusing potentially vulnerable consumers. A solution could be to let market forces and consumer preferences for common standard offers eliminate those firms that follow such tactics. Our literature review covers work that, like ours, deals with consumer confusion and firms' attempts to profit from it. We conclude that only an experimental approach will allow us to determine if consumer do indeed prefer common standards. We therefore present an original experimental setting that reproduces in a simplified manner the kind of issues consumers may face when choosing among products that are presented in different ways. We show that within that setting there is a range of different rules a consumer might follow when making choices. We describe the characteristics of the participants in our experiment and do some exploratory analysis of their decisions, which informs our last section where we deal with the econometric analysis of our data. Our conclusion summarizes our findings, underlines their consequences and offers ideas for further research.

1 Context

Behavioural economics finds that consumers have "inconsistent, context dependent preferences" and may not have "enough brainpower to evaluate and compare complicated products" (Spiegler, 2011). They often do not make the best choices for themselves, whether from an external, objective point of view or, more damaging, from their own point of view as well, that is, "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). Not only do they make what they would themselves consider bad choices, but those choices may not even make them happy, whether in the short or the long term . Normative economics asks what ought to be done about it, and differs in its advices. Three schools of thoughts are particularly salient: The libertarian paternalist approach (Camerer et al., 2003; Thaler and Sunstein, 2008) lists errors people typically make and suggests ways to get people to do what they would have anyway done *absent their limitations.* The welfarist approach (Layard, 2005) seeks to maximize consumers' experienced utility (happiness), possibly manipulating their choice, of leisure for example, through taxes. The liberal approach argues consumers ought to be free to choose as they wish and the market free to fulfill their needs as they occur (Sugden, 2004). The first school appeals to the concept of a consumer's "best self", the second invokes a consumer's greater good, and the third defends a consumer's sovereignty and freedom of choice.

What would a libertarian paternalist do about firms that try to confuse consumers with offers that are difficult to evaluate and compare with the competition?[2] An unvarnished paternalistic approach would be to get firms to make their offers easier to compare, or limit choice to those offers that have been determined to be best. There are issues with this approach: it is rather heavy handed and requires defining standards from the top down, which may result in too much rigidity and hinder further innovations in a sector. It also supposes one is confident that an expert may know better than the consumer, which is doubtful, see Freedman (2010) for example. A libertarian paternalistic approach would be to make sure consumers choose the best product without constraining their freedom to make bad decisions as well. Being paternalistic, it suffers from the same problems as above but, relying as

[2]We did not find recommendations on this topic from happiness researchers.

it does on subtle, “beneficial” framing of choice problems, the consumer may not even be aware of how his choice is being manipulated while “straight” paternalism at least does not disguise itself.

More satisfying, if one wants to involve a third party in the solution of the problem, would be 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). Better education in the basic principles of home economics, more particularly consumer science and resource management, could also be (re-)introduced.

Gaudeul and Sugden (2011), however, argue that all the above is at least in large part unnecessary as competition ought to drive firms to simplify their offerings *if at least some consumers follow the common standard rule*. Regulatory intervention would thus be limited to preventing firms from colluding to keep different formats for their offers. However, their conclusions hold only if indeed some consumers prefer offers that are expressed in ways that make them easy to compare with others. The purpose of this paper is to check if such consumers indeed exist, and if so, if their number is sufficient to drive firms into making their offers simpler to compare with others.

At this point, it is helpful to explain further why we concentrate on the study of the common standard rule. The CS rule is such that not only will consumers avoid the higher priced of the common standard offers, but they will choose the lower priced of the CS offers (the “LPCS” offer or “LPCS” for short) and disregard individuated standard (“IS”) offers. There are many ways in which we can justify this behavior:

1. *Statistically*, if one assumes that prices are i.i.d. across firms and firms choose whether to be CS or not at random, then the LPCS has lower prices in expectation than other firms. As in the Monty-Hall problem (Friedman, 1998), there is information gained from being told that an option is dominated.

2. *Behaviourally*, consumers (human or not, see Latty and Beekman, 2010) 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 (Busemeyer and Townsend, 1993). Note however that while the literature on this topic considers the dominated offer to be *irrelevant*, it is very much relevant in our setting as per the statistical argument above.

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. Consumers ought therefore to learn to behave optimally over time (Sugden, 1986; Fudenberg and Levine, 1998). This learning should be made easier by the general applicability of the common standard rule to

many environments, so that consumers who learned from one environments that CS firms have lower prices on average than other firms will apply this insight generally. Note that this does not require that consumers understand why this principle seems to be effective, and they may in fact rationalize it in other ways than the way we do so here.

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, even if in some situations it is not the best way to make choices (Sugden, 1989). 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 behaviour and heuristics, marketing, consumer protection and the competitive process.

2 Literature review

There has been a recent increase in interest within the field of industrial organisation for how consumers biases, limitations and inconsistencies can be exploited by firms (Carlin, 2009; Chioveanu and Zhou, 2009; Ellison, 2005; Gabaix and Laibson, 2006; Piccione and Spiegler, 2010). Ellison (2006) provides a survey of this literature and Spiegler (2011) wrote an excellent book analysing how boundedly rational consumers affect conclusions from conventional IO. Yet, Ellison (2006) observed that “in many papers in the IO branch, the particular form of the departures from rationality is motivated by little more than the author’s intuition that the departure is plausible” and suggested “more reliance on empirical and experimental evidence”. Recent research goes into evidence of how firms might design their offers to exploit consumers (DellaVigna and Malmendier, 2006; Ellison and Ellison, 2009; Miravete, 2003, 2011). Experiments are still quite rare however, though the fields of marketing (Morwitz et al., 1998; Viswanathan et al., 2005; Zeithaml, 1982) and psychology (Ariely, 2008; Iyengar and Lepper, 2000; Iyengar et al., 2004) have contributed to our understanding of how consumers deal with products choices in realistic purchasing scenarios. Economists have also led some experiments, among others Huck and Wallace (2010) and Shestakova (2011), who consider how different price frames affect consumer choice, and Kalaycı and Potters (2011), who consider how more complex offers increase firms profits in an experimental duopoly setting. We were however particularly inspired by field data on how functionally illiterate consumers make their choices (Viswanathan et al., 2005).

A common thread in this literature is that consumers tend to rely on simple heuristics to navigate complex tariff structures. As Ayal (2011) puts it: "reality shows that when complexity is high, consumers tend to seek decision tools [...] reducing the dimensionality of the problem, usually through heuristics limiting the number of attributes considered". Payne et al. (1993) suggest this is the result of an effort-accuracy tradeoff but Gigerenzer et al. (2000) argue that "less knowledge (can) lead to systematically better predictions than more knowledge" so that heuristic decision making can perform better than strategies that require more effort and information.

Our paper contributes to the existing literature on confused consumers by relying on a slightly framed laboratory setting in which values are induced, repeated observations are possible, the best choice is well defined and consumer mistakes and biases can be controlled for. Our setting allows us to define an exhaustive set of decision rules that subjects may follow when faced with decisions in our experiment. This would not be possible when using empirical data. Looking at product sales, for example, would introduce various confounds: the presence of real along with spurious product differentiation; regulations that may impose standards for a variety of reasons; economies of scale and network effects that may encourage the convergence to a technological standard; reputation concerns that may lead firms not to wish to confuse consumers; framing other than the standard adopted by the offer that may influence choice as well; habits and tastes such that the consumer chooses a product based on past purchasing behavior, and so on.

Doing an experiment in the laboratory allows us to study the demand side in isolation of the supply side, and makes it possible to create genuine spurious complexity, that is, complexity that all consumers will agree should be irrelevant to their choice. We can rule out regulation, economies of scale, network effects, reputation, habits and other influences on the subjects' choices. Since utility is well defined in our experiment, we can define precisely what is the best choice in each choice situation and exclude heterogeneity in consumer preferences. However, in order to keep the laboratory experience close to a purchasing act, the experiment is framed as a real buying decision, in which the participants are asked to *buy* a product out of *menus* of *offers* with the aim of minimising expenditure.

3 Experimental design

Participants faced two distinct set of tasks in our experiment. In the first part, the core of our experiment, the participants were faced with 80 different menus, each menu consisting of 3 to 6 options, each option being characterised by a set of attributes. Participants were asked to choose one option within each menu with the aim of minimising their expenditure. The second part was designed to assess participants' ability with respect to dimensions of relevance to the decisions they had to make in the first part. The participants were asked to tackle some (non incentivised) standard tasks to assess their ability to compare shape sizes, solve basic binary operations and deal with simple consumption decisions. The experimental software, the menu generator and the script to collect and organise the raw data were programmed in Python (van Rossum and Drake, 2001).[3] The English translation of the original German instructions is available in appendix B.

3.1 Purchasing task

The main task consisted of a purchase decision. Consumers were given a budget B and presented with menus of options, each option being described by its size, shape and price. More specifically, the task of the participants was to buy *paint* to cover a fixed, square area, and they were presented with offers consisting of a price p and a visual representation of the area that could be covered for that price. Participants were told that paint quality did not differ across offers.

Denoting s the size of the area covered by the offer as a proportion of the total area to be covered (s was always less than 1) their total expense was calculated automatically as p/s. Their payoff was what they managed to save from their budget B once the paint had been bought, that is, $B - p/s$; in payoff terms, all the participants had to do was to minimise their expenditure. The area covered by the offer was presented in different shapes and sizes, the {shape, size} combination being what we will henceforth call a *standard*. A common standard in this setting was an offer that has an equivalent in terms of size and shape in the menu. In that case, the consumer needed only choose, among the options that adhere to the common standard, the one that minimizes p.

Presenting offers in terms of a combination of a shape and a size conveys the idea of a standard while drawing on an existing body of research on shape perceptions (Krider et al., 2001). It is easier to compare an offer with another if they have the same shape. If two items have the same shape and the same size, comparison is even easier.

Offers were thus described by three elements:

1. A shape out of three possible options: circle, square, equilateral triangle (see Figure 1). We considered only three shapes to limit confusion and to be able to build on the existing literature on shape comparisons, which is limited essentially to those basic shapes (Krider et al., 2001). There was no variations in terms of the positioning of the shape – always in the center of the background square that represented the total area to cover, resting on a base in the case of the triangle, resting on a side in the case of the square.

[3] 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 analysers, user interface) and its documentation, as well as the raw data and the script used to collect and organise them are available upon request.

2. A size out of 12 possible options. Normalising to 100 the size of the square area that must be covered, these options ranged in size from 10 to 43, in steps of 3 (see Figure 2).[4] The step was chosen to be big enough to allow for easy comparison within menus (at least if areas were of the same shape) while being small enough to have a sufficient number of steps to rule out learning and comparisons across menus.

3. A price, indicating how much it costs to cover the area covered by the offer. Unit prices ("up", the cost to cover an area of size 1 as per the normalization above) were drawn from a uniform distribution of mean 0.5 and standard deviation σ^2 equal to either 0.05 or 0.01. Standard deviation of 0.05 was meant to represent a case where there is a low degree of competition between products, while standard deviation of 0.01 represented a case with strong competition. The price of an offer, which was the information conveyed to the participants, was obtained from the unit price by multiplying it by the size of area covered by the offer.

Figure 1: The shapes used in the menus

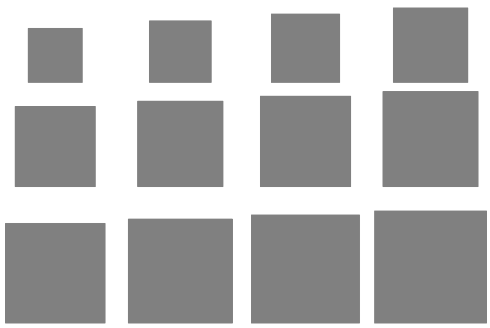


Figure 2: The sizes used in the menus

The offers were displayed as a grey area on a white background. The participants were told that, once they chose an offer, they would obtain for the shown price an amount of paint suitable to cover an area equivalent to the shown area. The white background was the total area to be covered. This allowed participants to visually appreciate the size of the shape with respect to the total area to be covered. The background was overlaid with a grid of thin light blue lines to ease comparison between offers.

Menus varied in terms of length (*3* or *6* offers per menu) and in terms of strength of competition between offers (*high* or *low*) in order to obtain a two-by-two within-subjects

[4]The size was limited to 43 as an equilateral triangle resting on a base cannot cover more than $5 \times \sqrt{75} =$ 43.301...

treatment structure. An example of a menu with three elements and a common standard (the triangle) is shown in figure 3. An example of a menu with six elements and no common standard is shown in Figure 4.

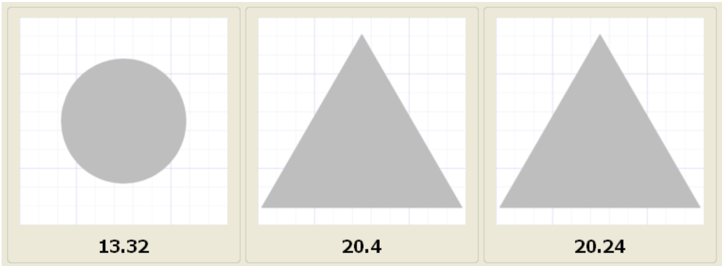


Figure 3: Screenshot of a menu with three offers and a common standard

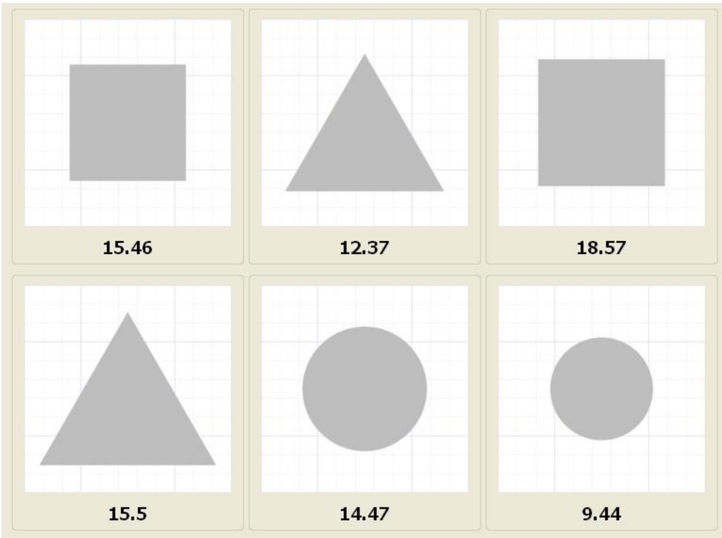


Figure 4: Screenshot of a menu with six offers and no common standard

Menus were randomly generated with Python under the constraint that no offer was to give, if chosen, a negative payoff to the participant. Moreover, within this structure, menus had either no common standard, one common standard or two common standards:

1. Menus with *no* common standards were such that a given {shape, size} combination would appear only once within the menu. Note that this requirements does not prevent the same *shape* from occuring more than once in a given menu, only that a *standard*, that is a {shape, size} combination, does not appear twice. It is therefore possible that menus with no common standard may present what we call *soft standards*, that is, menus with two offers of the same shape but different sizes.

2. Menus with *one* common standard, such that two (and only two) offers featured the same {shape, size} combination. Those were constrained not to have the same price.

3. Menus with *two* common standards (only possible for menus of six options), whereby one {shape, size} combination occurred twice while another occurred thrice.

Each individual was faced with 80 menus. 36 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 with two members, the other with three). In each case, half of the menus featured strong competition (variance 0.01), the other half weak competition (variance 0.05). The distribution of menus is summarized in table 1 below.[5]

Table 1: Distribution of menus by CS and strength of competition

		Strong competition	Weak competition
3-menu	No CS	9	9
	One CS	9	9
6-menu	No CS	9	9
	One CS	9	9
	Two CS	4	4

Once an option was chosen, the computer automatically purchased the exact quantity needed to paint the background square area at the unit price implied by the offer. If the participant chose an offer of size s and price p, which implied a unit price $up = p/s/100$ then the computer calculated her expenditure of $100 \cdot up$ and her payoff $B - 100 \cdot up$. The participant was thus best off choosing the option that minimised up.

The participants had up to two minutes to choose an offer from each menu. The choice was performed by clicking on the shapes - in which case they would be highlighted with a light green frame - but could be revised as many times as one wanted, within the two minutes limit, and was finalised only by clicking on a 'Submit' button present at the bottom of the screen (see appendix C). Finalizing a choice could not be done before 10 seconds had elapsed. If no final choice was made within the two minutes limit - that is, the participant had not clicked on the Submit button - the last highlighted offer was submitted as the final choice; if no offer was highlighted, then the participant would receive a payment of 3 euros for that trial, which is less than the minimum value of any offer across all our menus.[6]

The participants were given detailed feedback about their payoff in each period (see Figure 5). The feedback reminded them of their choice, and told them how much they spent to cover the required area. Their payoff was communicated as endowment minus expenditure.

[5]The menus are available on request for visual inspection. Apart from the visual representation provided there, the menus are described by lists of three or six 6-elements vectors (shape, size, unit price, shown price, dummy for soft standard, dummy for standard). These are available in comma separated files bundled with the experimental software, and can be sent upon request.

[6]Only one participant failed to make a decision within the time limit, and this only once, in that case providing no choice.

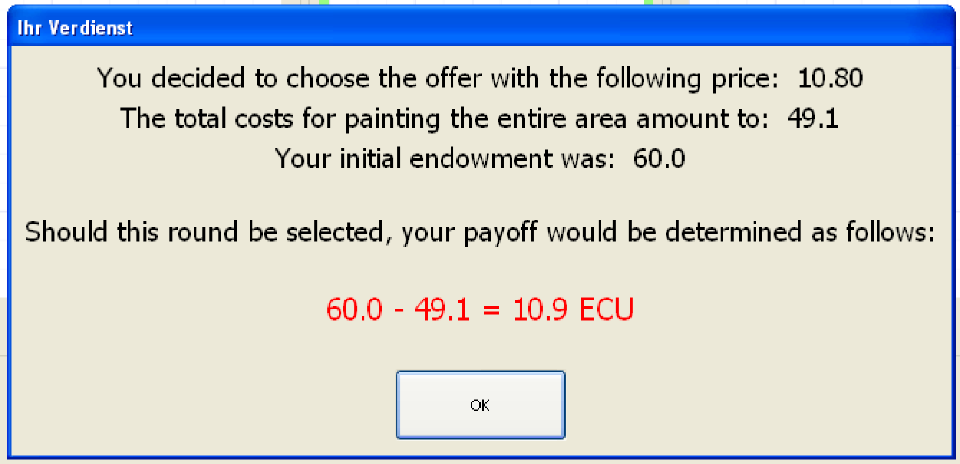


Figure 5: Feedback after each choice

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 many did record their payoffs. Once they confirmed the feedback by clicking on 'OK', they would be shown the next menu. The participants knew the total number of menus was 80 and reminded of their progress along the experiment.

3.2 Control tasks and questionnaire

The second part of the experiment consisted of non-incentivised tasks to control for the ability of the participants in performing the main task (see appendix D). Those tasks relate to visual perception and computational skills. Three different set of tasks were chosen:

1. *Shape comparisons*: The participants were asked to give their estimate of the relative size of a shape with respect to another. The comparisons involved rectangles, circles and triangles. Each comparison had to be done within a time limit of one minute. No minimum time was enforced and the participants could skip any question.[7]

2. *Mathematical operations*. The participants were asked to solve 3 sets of 10 operations (sum, subtraction, multiplication, divisions). Each set had to be completed within one minute. No minimum time was enforced and the participants could skip a set or not fill in the result of an operation. The sets were generated using Mail Goggles's GMail Labs app by Jon Perlow[8] and were graded in terms of difficulty.

3. *Consumer problems*: The participants were asked to solve three simple problems that were expressed in simple terms. The first tested their understanding of the concept of area, the second and third were designed to test whether they understood how an area relates to its dimensions, and the fourth checked whether they were able to translate

[7] Only one participant did so.
[8] http://gmailblog.blogspot.com/2008/10/new-in-labs-stop-sending-mail-you-later.html

a number from one standard to another (here, a currency). Each problem had to be solved within one minute. No minimum time was enforced and the participants could skip any problem.

Once done with the control tasks, the participants filled in a short questionnaire designed to assess their demographic profile. 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 drew from a urn a number from 1 to 80, and was paid according to her purchasing decision in the drawn period.

4 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 choice and then take into account other factors in our empirical model. We consider two 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. 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.[9] From those two criteria, we can derive four possible heuristics, illustrated in the following graph and explained below.

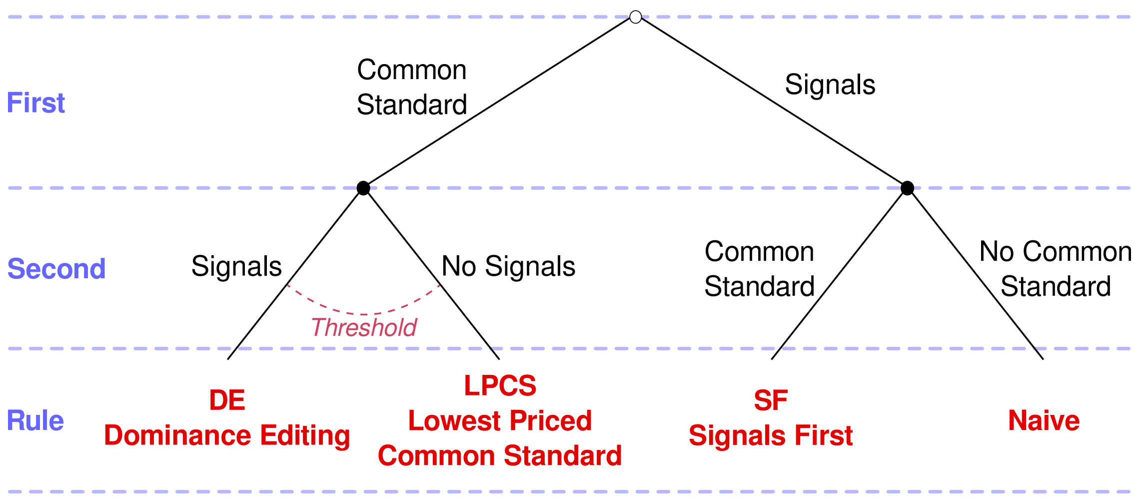


Figure 6: Choice criteria and heuristics

[9]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.

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 editing ("DE"). On the right, if the consumer first considers signals, he may provisionally choose the firm 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 Naïve 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, reverts to the Naïve rule otherwise, the DE rule determines $k = \arg\min_{i \in CS} p_i$ if there is a CS, in which case it chooses $\arg\min_{i \notin CS} (\hat{up}_{kj}, \hat{up}_{ij})$, reverts to the Naïve 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.

As evoked above, many other rules may be followed, among those some we label as follows:

- The *budget rule* chooses arg min p_i. This is a rule that favors small packages. Viswanathan et al. (2005) says it may be followed by those consumers who face budget constraints that prevent them from spending more than a fixed amount. This does not make sense in our setting, except maybe for another reason, which is that it may be easier to evaluate how many times an area is contained into another if that area is small relative to the other. Smaller areas would thus be chosen because they provide more certainty as to how much will be spent overall when choosing them.

- The *bulk purchasing rule*, whereby some consumers may favor big packages, which is justified if one considers that the cost of the material in a package decreases in relation to the value of its content as its size increases, so offers in big packages are usually better deals than those in small packages.

- The *omniscient rule* chooses $\arg\min_i up_i$, and corresponds to the Naive rule with no error term. This may be followed if a consumer spends considerable time measuring the area of each offer and then computes the ratio p/s for each offer so as to determine the best offer. This may be the case of poor but numerate consumers (Viswanathan et al., 2005).

- The *random rule* chooses offers at random, and corresponds to the Naive rule when e_{ij} is very large.

- *Lexicographic rules* 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. They may also favor the last offers in the menu, if the consumer tends to remember (and choose) the last option they read from a list.

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

4.1 The Threshold rule

The DE and LPCS rule are in fact only extreme cases in a larger category of what we call Threshold rules that 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 v_j depending on consumer j's preference for ($v_j > 1$) or against ($v_j < 1$) the LPCS. The optimal choice of 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. v_j=1 corresponds to the DE rule, $v_j \to \infty$ 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 scepticism 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 even clearer, suppose the consumer compares two oranges and one apple, with oranges being priced at \$0.45 and \$0.55 respectively, while the price of the apple is \$0.65. The consumer cares only about calories, of which he needs 2000 per day for a work paid \$30 per day. His utility is expressed as $30 - p \times 2000$ with p the price per calories of his food. He *estimates* the oranges to contain 35 calories each, while the apple *appears to* contain 55 calories. He compares the lower priced orange with the apple in terms of price per calories. The lower priced orange appears to cost \$1.29 per 100 calories, while the apple appears to costs \$1.18 per 100 calories. The consumer will however choose the apple only if his threshold is less than 1.09. This makes sense if for example he is not sure about the respective calorie content of oranges *vs.* apple but is sure that both oranges have the same number of calories.

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 our context where offers are randomly generated, so that it is always better for a consumer to follow the Threshold rule with $v_j > 1$. 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.

4.2 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 Na works best and LPCS is worst. Indeed, taking an example with three menus for ease of notation $B - E(\min(a, b, c)) > B - E(\min(a, b))$ with B the budget and a, b, c i.i.d. random variables. On the other hand, if consumers make considerable mistakes, then Na is worst (it results in choosing essentially at random) while LPCS is best. Indeed $B - E(\min(a, b)) > B - E(a)$. We performed simulations with Octave (Eaton, 2002) to examine the performance of each rule in terms of expected consumer payoff.[10] 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 (case with strong competition), and 0.05 (case with weak competition) and B was set to 60. Consumer choice was simulated according to the various rules expressed above (Na, DE, LPCS, SF), as well as according to the Threshold rule, with the optimal threshold 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. Note however that the ranking of payoffs by rules is quite robust and differences in payoffs by rule are significant for low number of choices.[11]

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 prices are drawn from a distribution corresponding to either weak or strong competition. Also shown on separate scale is the optimal threshold v^* for each value of the error term.

[10] Program available upon request.
[11] In all the following, “significant” will be understood to mean “significant at the 5% level at least”.

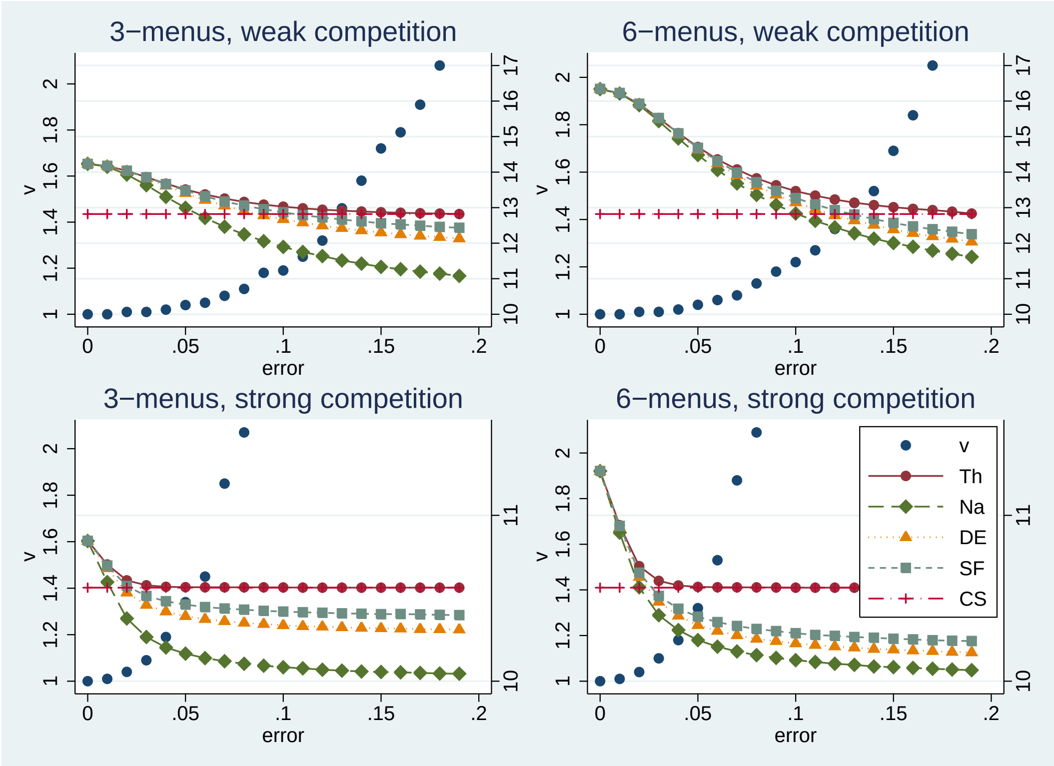


Figure 7: Consumer payoffs by choice rules and optimal thresholds, by menu length and strength of competition

As can be seen, payoff decreases as consumers become less accurate in their choice, except for the LPCS 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 higher levels in σ^2. Following the CS rule obtains higher payoffs than the DE, SF or Naive rules as long as σ^2 is not too low. The CS rule is better than those other rules even for rather precise consumers when competition is strong, as even high levels of accuracy may result in mistakes if prices are close together. In terms of ranking of rules, the Threshold rule outperforms the CS rule, while both SF and DE dominate Na, 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, DE saves time and effort compared to the SF because it requires estimating the value of a lower number of alternatives than the SF.

5 Descriptive statistics and exploratory data analysis

We ran our experiment with 202 subjects at the Max Planck Institute in Jena on June 10, 14 and 15 in 2011, over 8 sessions with 24 to 27 subjects each. 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, with 75 studying social sciences, of which 21 studying economics and 10 business administration. The rest were studying languages (21), natural sciences (16), or were studying to become teachers (24), jurists (15) or health care professionals (13). 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 the shape comparison task	0.25	0.25	0.10	0.35	0.05	0.58	201
Score in practical problems	2.78	3.00	0.96	-0.27	1.00	4.00	202
Score in mathematical problems	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
Mean individual pay-off	11.44	11.48	0.41	-0.80	9.88	12.28	202
Mean time spent on each task	19.67	18.34	6.36	1.30	11.66	46.27	202

The average age of our subjects was 23.65, 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 6.23, 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). The fastest subject took 12 seconds on average on each menu, meaning he spent 16 minutes on this part of the experiment, while the slowest took 46 seconds per menu on average, thus spending a bit more than one hour on the purchasing tasks.

As for the control tasks, in the test of the ability to assess the area of one shape in terms of multiples of another, we computed for each individual the average deviation from the true value of the ratio of the size of one shape to the other 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 57%. In the task to test mathematical ability, 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. Subjects performed less well with practical consumption problems with only about 62% answering more than half of the questions correctly. Performance in the different tasks were significantly and positively correlated, though not highly (correlation coefficients were around 0.35). Women performed less well than men in all tasks.

5.1 Individual payoffs

Let us now consider individual payoffs by menu length, strength of competition and presence of a CS (Table 3).

Table 3: Payoffs by menu length, strength of competition and CS

		Strong competition			Weak competition		
		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 4 thus reports the average of the ratio $(up^{max} - up^{chosen})/(up^{max} - up^{min})$ over individuals and tasks in each category. We will 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 4: Performance ratio by menu length, strength of competition and CS

		Strong competition			Weak competition		
		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 been careful in their choices. As could be expected from statistical arguments, individuals obtained higher payoffs with 6-menus and when competition was less intense.

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 6-menus with one CS and strong competition. The presence of a CS did not consistently improve consumer payoffs when competition was strong, but significantly and consistently increased payoffs when competition was weak.

Regressions of payoffs on individual and menu characteristics indicate that women obtained better payoffs,[12] payoff increased with the order in which the task was presented (so there was some learning) and subjects with higher scores in the mathematical and practical tasks obtained higher payoffs as well. Motivation, scores in the shape comparison task and time spent on a task did not appear to have a significant effect.[13] 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 task scores. Lower strength of competition, longer length of the menu and the presence of a CS also increased payoffs. The effects above are robust to various specifications.

Overall, consumers made about 39% of their choices correctly, that is, choosing the firm

[12] We will see later that this might be due to their better use of CS information.

[13] 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.

with the lower 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. 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.

When mapping payoffs by menu length and strength of competition (rows with no CS in table 3) to the predictions from our simulations for those menus with no common standard (Graph 7), 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 competition is strong. Lower accuracy when choosing within menus where competition is weak did not prevent them from obtaining higher payoffs than when choosing within menus where competition is strong however. 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. Consumers did not obtain higher payoffs in 6-menus than in 3-menus when competition was strong, 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 consumers generally chose thresholds that were lower than this, indicating perhaps that they were over-confident in their ability to make accurate choices.

5.2 Firms sales and profits

Table 5 shows that the LPCS was chosen about 56% of the time within our 3-menus, a lot more often than the IS. Similarly, the LPCS was chosen about 25% of the time in 6-menus with only one CS, while each of the four IS offers were chosen only about 18% of the time. We do not report results for 6-menus with two CS from this point on.[14] From this, we can conclude that consumer are not all following the DE rule, as this would have resulted in choice probabilities being divided equally across all offers except the HPCS. They are not all following the CS rule either, since the LPCS would then have been chosen with probability one. Their choices would be more consistent with the SF rule overall since consumers appear to divert their choices from the HPCS to the LPCS without greatly affecting the probability with which a IS is chosen. We will see later, however, that this is an artefact of the aggregation of individual choices, as individuals follow a mix of rules and few actually follow the SF rule.

[14] This is both for ease of exposition and because results with two CS would require that we extend further our tables to include the LPCS and HPCS of the larger standard and of the smaller standard.

Table 5: Choice frequencies by menu length, strength of competition and CS

		Strong competition			Weak competition		
		LPCS	HPCS	IS firm	LPCS	HPCS	IS firm
3-menu	No CS	NA	NA	33.33%	NA	NA	33.33%
	One CS	53.63%	7.21%	39.11%	59.79%	4.51%	35.70%
6-menu	No CS	NA	NA	16.67%	NA	NA	16.67%
	One CS	25.63%	4.07%	17.58%	25.47%	2.97%	17.89%

Note: There was one instance in which a subject did not make a choice in time, so the sales do not add up to 100% in the case of the strong competition 3-menu with a CS. In 6-menus, sales for an individual IS firm are calculated by averaging sales by IS firms.

Consider an *hypothetical* firm that would have been the one making the LPCS offers within our menus, and let us call it a "LPCS firm". Its profit equals revenues since we do not consider production costs. That LPCS firm would have sold at a lower price than either IS or HPCS firms on average, but its profits would still be higher than either IS or HPCS profits since the LPCS was chosen more often (Table 6). More precisely, profit of an LPCS firm would be 0.27 on average when there are three options, compared with profit of 0.17 for an IS firm firm, and 0.12 on average when there are six options and one CS, compared with profit of 0.09 for an IS firm.[15] Note however that an IS firm would still make higher profit when there is a CS than when not as the number of competing firms decreases with the ensuing quasi-elimination of the HPCS firm.

		Strong competition			Weak competition		
		LPCS	HPCS	IS firm	LPCS	HPCS	IS firm
3-menu	No CS	NA	NA	0.1653	NA	NA	0.1632
	One CS	0.2640	0.0362	0.1949	0.2811	0.0233	0.1622
6-menu	No CS	NA	NA	0.0831	NA	NA	0.0800
	One CS	0.1271	0.0205	0.0880	0.1110	0.0147	0.0840

Table 6: Profit by menu length, strength of competition and CS

Note: In 6-menus, profits for an individual IS firm are calculated by averaging profits by IS firms.

In summary, an LPCS firm would make higher profits than other firms, consumers obtain higher payoffs when there is a CS, and the price at which LPCS offers are sold is lower than the price at which other offers (HPCS, IS) are sold. There would therefore be an incentive for an IS firm to switch to a CS, and also an incentive for consumers to choose the LPCS offers. While this may not necessarily translate into a convergence to a CS as hypothesized in Gaudeul and Sugden (2011) (this will be tested in a further article), the conditions are in place for this to be so.

[15] One can retrieve average sales price by the formula profit=sales×price, using data in tables 5 and 6.

6 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.

6.1 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 where 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 price of the option, its shape and its size. Shape is coded from most attractive to least attractive, which means that a triangle is assigned a value of 1, a square a value of 2 and a circle a value of 3 as shapes that extend more broadly in space are preferred (see Krider et al., 2001). If an alternative specific constant was included in the model and was significant for some options, this would mean that position in the menu influences choice. 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 the experiment." We however consider a variable "position" which takes values 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. This allows us to determine if lexicographic position in the menu influenced consumer choice.

Formally, denote y^o_{ijm} the utility of option j in menu m for invidual 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 $y^o_{ijm} = \alpha\, up_{jm} + \beta\,\text{shape}_{jm} + \gamma\,\text{area}_{jm} + \phi\,\text{position}_{jm} + u_{ijm}$ with u_{ijm} a random variable of mean 0 that follows either a logistic or a normal distribution.

We find that a full model that takes into account all the above variables minimizes the Akaike Information Criterion ("AIC"). Results are shown in table 7 and indicate that a multinomial probit model is preferred to a logit model, so error terms are best seen as following a normal distribution as was done in our simulations. 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. Results from the mixed logit model indicate there is significant variation in the extent to which irrelevant factors (shape and size) influenced consumers.

6.2 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 variable (here, a case is an individual) along with alternative-specific variables to determine choice among alternatives. Our case specific variables are individual scores in the mathematical, shape comparison and practical problems, along with gender, time spent on a task and motivation.

The model above is thus modified as follows: Latent utility y^o_{ijm} takes the form $y^o_{ijm} = \Omega_i\lambda_j + \alpha\, up_{jm} + \beta\,\text{shape}_{jm} + \gamma\,\text{area}_{jm} + \phi\,\text{position}_{jm} + u_{ijm}$. As before, j is the option, m is the menu and i is the invidual. An option is coded in terms of whether it is the LPCS, the HPCS or an IS offer, in which case it is coded as IS1 to IS4 depending on its position in the menu. Ω_i is a $1 \times q$ vector of case-specific variables, the same variables being assumed to influence the choice for each option, and λ_j is a $q \times 1$ 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. u_{ijm} is a random variable of mean 0 that follows either a logistic or a normal distribution. We constrain λ_j to be the same for all 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, ought to be used as case-specific variables. Results are reported in table 6.2.

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). Consumers tend to favor op-

Table 7: 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						
up_	−14.1925∗∗∗	−13.1827∗∗∗	−15.7299∗∗∗	−14.3147∗∗∗	−6.1562∗∗∗	−14.8987∗∗∗
	(−13.72)	(−12.29)	(−14.69)	(−20.56)	(−6.98)	(−18.72)
pos_	0.0453+	−0.1073∗	0.0443+	0.0071	0.0467∗∗	0.0065
	(1.84)	(−2.19)	(1.82)	(0.71)	(3.19)	(0.62)
shape_	−0.3623∗∗∗	−0.3588∗∗∗	−0.3981∗∗∗	−0.3106∗∗∗	−0.1279∗∗∗	−0.3724∗∗∗
	(−12.39)	(−11.78)	(−9.42)	(−13.86)	(−6.19)	(−9.22)
ar_	−0.0137∗∗∗	−0.0150∗∗∗	−0.0154∗∗∗	−0.0049∗	−0.0022∗∗	−0.0052
	(−6.07)	(−6.09)	(−4.45)	(−2.47)	(−2.64)	(−1.20)
SD						
shape_			0.3869∗∗∗			0.4469∗∗∗
			(10.14)			(9.79)
ar_			0.0355∗∗∗			0.0530∗∗∗
			(8.27)			(12.22)
N	10908	10908	10908	21816	21816	21816
ll	−3785.9059	−3776.8929	−3715.3185	−6162.2205	−6095.4068	−5944.5553

t statistics in parentheses

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

Table 8: 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
main						
up_	−15.0358 * **	−11.8822 * **	−16.1881 * **	−13.5463 * **	−5.4703 * **	−13.8677 * **
	(−13.80)	(−5.18)	(−13.30)	(−23.94)	(−11.23)	(−21.44)
pos_	−0.0314	−0.0335	−0.0220	0.0606 * **	0.0382 * **	0.0562 * **
	(−1.07)	(−1.41)	(−0.75)	(5.72)	(5.51)	(5.25)
shape_	−0.1722 * **	−0.1399 * **	−0.1800 * **	−0.3299 * **	−0.1315 * **	−0.3917 * **
	(−5.04)	(−3.58)	(−4.04)	(−13.87)	(−9.77)	(−9.54)
ar_	−0.0062*	−0.0048+	−0.0040	−0.0081 * **	−0.0031 * **	−0.0085*
	(−2.05)	(−1.80)	(−1.02)	(−4.57)	(−3.98)	(−2.32)
HPCS						
scoreshtask	1.8358*	1.4086 * *	1.8457+	3.2298 * *	1.2695 * **	3.2550 * *
	(2.43)	(2.66)	(1.91)	(2.97)	(3.78)	(2.86)
scoremath	−0.0098	−0.0109	−0.0093	0.0675+	0.0293*	0.0688+
	(−0.42)	(−0.66)	(−0.31)	(1.87)	(2.30)	(1.84)
genre	−0.4795 * *	−0.3000 * *	−0.4811*	−0.8517 * **	−0.3129 * **	−0.8538 * **
	(−3.23)	(−2.85)	(−2.43)	(−4.31)	(−4.29)	(−3.66)
_cons	−1.8650 * *	−1.7919 * **	−1.8470*	−3.3464 * **	−2.0394 * **	−3.3768 * **
	(−3.20)	(−4.25)	(−2.49)	(−3.68)	(−5.95)	(−3.62)
IS1						
scoreshtask	−0.3171	−0.3814	−0.4251	0.3797	0.2356	0.3670
	(−0.82)	(−1.16)	(−0.98)	(0.87)	(1.32)	(0.83)
scoremath	0.0245+	0.0215+	0.0279	0.0054	0.0016	0.0060
	(1.82)	(1.83)	(1.53)	(0.37)	(0.26)	(0.35)
genre	−0.2633 * **	−0.1887*	−0.3087 * *	−0.2511 * *	−0.1187 * *	−0.2655 * *
	(−3.42)	(−2.54)	(−3.20)	(−2.80)	(−3.10)	(−2.71)
_cons	−0.6191+	−0.6055*	−0.6825	−0.1603	−0.0904	−0.2033
	(−1.85)	(−2.02)	(−1.53)	(−0.44)	(−0.55)	(−0.47)
SD						
shape_			0.3541 * **			0.4513 * **
			(7.25)			(10.77)
ar_			0.0367 * **			0.0437 * **
			(8.42)			(11.78)
N	10851	10851	10851	21708	21708	21708
ll	−2948.2200	−2944.5503	−2880.4806	−5647.8391	−5587.0351	−5480.4975

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The number of observation is lower than in regressions with a CS because one subject did not perform the shape comparison tasks. Base outcome is the LPCS.

tions at the end of the 6-menus, maybe because of difficulties in recalling options that are previous in terms of their lexicographic order. In terms of case-specific variables, consumers tend to avoid the HPCS: the parameter on the constant term for that option is negative and highly significant. However, it is only in the case of 3-menus that consumers display an aversion to the IS *vs.* the LPCS, while this is not the case in 6-menus. Individual 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. Interestingly, those subjects that are better at the mathematical task tend also not to avoid the HPCS as much as others, though the effect is small and appears only for 6 menus. Women seem to be better than men at avoiding the HPCS, and also seem to avoid the IS, thus being closer to following the LPCS rule. In conclusion, all consumers tend to avoid the IS in 3-menus, though this is more pronounced for women, while only women appear to keep on following the LPCS rule when choosing among 6-menus. This might explain why women managed to obtain higher payoffs than men in this experiment even though they were less good at tasks that predict higher payoffs in our regressions.

A reason why some subjects would not display aversion to the IS in 6-menus while they do so in 3-menus may be found in the difficulty in identifying the presence of a CS in those menus. An indication of this, with reference to table 5, is that sales of the HPCS were higher in six menus than in 3 menus, which means more subjects did not realize there was a CS.[16] Following the CS rule in 6-menus might thus require that one have a more distinct understanding of it than in 3-menus since one has to be actively checking if there is a CS in 6-menus while the presence of a CS is almost self-evident in 3-menus.

6.3 Assigment 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 4. Consider table 9 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.

[16] Note that to translate sales from the 3-menus to the 6-menus, one has to divide them by two.

Table 9: 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^{Na}_{HPCS}	p^{Na}_{LPCS}	p^{Na}_{IS}
CS	0	1	0
SF	0	$p^{Na}_{HPCS} + p^{Na}_{LPCS}$	p^{Na}_{IS}
DE	0	$p^{Na}_{LPCS}(LPCS, IS)$	$1 - p^{DE}_{LPCS}$
Threshold rule	0	$p^{Na}_{LPCS}(LPCS, IS \times v)$	$1 - p^{Th}_{LPCS}$

In the case shown in table 9, 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 LPCS 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 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 one can exclude the HPCS from the consideration set, for example 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.

In terms of notations, we denote the probability the LPCS is chosen under the DE rule as $p^{DE}_{LPCS} = p^{Na}_{LPCS}(LPCS, IS)$. 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 $p^{Th}_{LPCS} = p^{Na}_{LPCS}(LPCS, IS \cdot v)$.

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.

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 thus 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.

We still face a problem in 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 (this one would tend to infinity). 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 in play. We thus compute for each consumer, the parameter 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.

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.

Comparison between rules will thus be done using the Akaike Information Criterion ("AIC"), but also, for the purpose of comparison, and even though we know this is not a correct way to estimate the performance of a model, by counting the number of choices that were "correctly" predicted by each rule, with "correct prediction" being taken to mean that the option that had the highest probability to be chosen according to one rule was indeed chosen.

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.

Using the same notations, the number of "correctly predicted" choices is $g(y,\theta) = \sum_{t=1}^{N} \mathbb{1}\,[y_{tj} p_{tj} = \max_j p_{tj}]$.

6.3.1 What rule best describes aggregate behavior?

Table 10 reports the log-likelihood, the values of the AIC and of the Bayesian information criterion ("BIC") and the sum of correctly predicted choices ("CPC") for each rule, for 3 and 6-menus. The last column contains the value of 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 LPCSs. For more interpretation of this number, see the detailed explanation in section 4.1.

Table 10: Rules scores, aggregate behavior

		Naive	Signal First	Dominance Editing	Threshold Heuristic	v
3-menus	LL	−3520	−2993	−3078	−2978	1.07
	df	6	7	7	8	
	AIC	7052	6000	6170	5972	
	BIC	7089	6043	6213	6022	
	CPC	2104	2148	2143	2209	1.09
	N	3636	3636	3636	3636	
6-menus	LL	−5953	−5727	−5734	−5703	1.05
	df	6	7	7	8	
	AIC	11918	11468	11482	11422	
	BIC	11955	11511	11525	11472	
	CPC	1137	1190	1277	1333	1.12
	N	3636	3636	3636	3636	

The choices in 6-menus appear to be considerably less accurately predicted under any type of rule than in the 3-menus. This means there is more randomness in consumer choice in 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. 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 5 to 7% 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 and attempt to determine rules followed by individuals in the next section.

6.3.2 What rules do individuals follow?

The above techniques were used to determine rules followed by the subjects. Table 11 cross-tabulates the number of subjects assigned to each type when looking at 3-menus and at 6-menus, summarizes our findings:

Table 11: Subjects assigned to rules, by menu length.

		6 menus				
		Naive	SF	DE	Th	Total
3-menus	Naive	26	7	6	1	40
	SF	25	23	21	3	72
	DE	17	15	22	2	56
	Th	9	11	10	3	33
	Total	77	56	59	9	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 (77 *vs.* 40 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 (56 *vs.* 72 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 4.2. 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 29 subjects of the 33 assigned to the Threshold rule that do favor CS in 3-menus, and the 6 subjects of the 9 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.

While there are no significant differences in the rules followed by different genders when faced with 3-menus, women are significantly more likely to follow the SF and significantly less likely to follow the DE rule when faced with 6-menus. As seen in this paper, the SF rule theoretically obtains higher payoffs than the DE rule, however, as seen above, this is not the case in our sample.

6.3.3 Do consumers that follow the Threshold rule 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 cases are more accurate. The following graphs relate average payoffs obtained by subjects in 3 and 6 menus with no CS to the threshold assigned to them by the procedure above. Bigger points indicate those individuals that were assigned to the Threshold rule. We superimpose on this graph the optimal choice of threshold for a consumer of the accuracy implied by his average payoff when faced with menus with no CS. 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 4.1, expectation in the formula determining v_j^* is taken over all menus of a specific length.

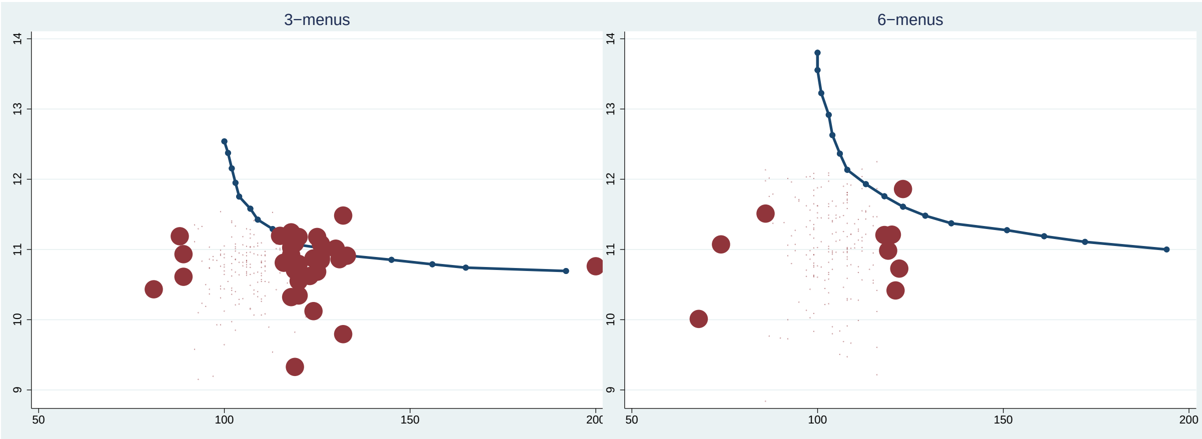


Figure 8: Optimal vs. realized threshold in 3 and 6-menus

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. This means that while consumers that use the Threshold rule can be considered as being more savvy than those who follow other rules, they are however 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. 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.

6.3.4 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 below:

Table 12: 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, 69% of individuals are classified the same way in 3-menus, and 74% 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 use the Threshold rule in later stages of the experiment, but the increase is small. 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. It seems however that subjects learn to realize that there may be common standards in 6-menus, where CS are less obvious. Indeed, fewer subjects appear to be naive when excluding the first 20 tries than when considering the whole sample.

7 Conclusion

We found that as many as 16% of consumers followed the Threshold rule when choosing among 3-menus, but that even with with some learning, less than 10% followed that rule when choosing among 6-menus. Those consumers who followed the Threshold rule applied a large price penalty to offers that were expressed in terms of an individuated standard however: An offer expressed in terms of an individuated standard had to be about 6% less expensive than an offer expressed in terms of a common standard before it would have the same likelihood to be chosen. In other terms, an offer expressed in terms of a common standard and priced at $1.07 is *as likely to be chosen* by one of our savvy consumers than the *same* offer expressed in terms of an individuated standard and 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. While this is a positive result, it is tempered by the fact that consumers did not seem to learn to follow the Threshold rule over the course of the experiment, and even those who adopted the Threshold rule adopted thresholds that were lower than would have been optimal given their inability to make accurate choices among offers.

Even though relatively few consumers followed the Threshold rule, and even when they did so adopted lower thresholds than optimal, consumers' aggregate behavior favored offers that were expressed in terms of a common standard. That this would be so even though many consumers followed other rules underlines the robustness of the Threshold rule and its ability to drive firms to adopt common standards. This disproportionate influence of savvy consumers is due to the Signal-First and Naive rule being neutral with respect to the existence of a common standard. Furthermore, while the dominance editing rule does not favor common standard offers *vs.* individuated standards, it still results in the lower priced common standard offer gaining market shares at the expense of the higher priced common standard offers, making the adoption of a common standard a profitable strategy. This confirms our belief that a process of convergence towards a common standard equilibrium should occur under a broad set of initial conditions (Gaudeul and Sugden, 2007).

We also saw that "too much choice" could work towards negating the common standard effect, in the sense that it made it difficult for consumers to identify offers that were expressed in terms of a common standard. While the Threshold rule may be effective in fighting against the introduction of *spurious complexity* by firms wishing to confuse consumers, it may therefore not be so 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. This means that, for a common standard effect to be effective in markets where there is a multiplicity of choices, firms ought 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 (Adams et al., 1997, p.93, point 7). 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 increasing its salt content. This means that if we accept that the effectiveness of the common standard effect is enhanced by firms being able to advertise their use of a common standard, then there may be a role for a regulatory authority that would promote and monitor the use of standards and mandate the disclosure of the information that enters into the definition of that standard.

In future work, we would like to test under what conditions firms in a competitive market may converge towards the use of a common standard. From the present paper, consumers that follow the Threshold rule have positive externalities on other consumers since they make it unprofitable not to adhere to a common standard. Whether this translates to a more complex competitive setting, and if so, subject to what restrictions, is an open question, but this paper will help us in simulating the market conditions on the consumer side.

References

Adams, A., A. Di Benedetto, and R. Chandran (1997). Can you reduce your package size without damaging sales? *Long Range Planning 24*(4), 86–96. (ref. p. 35).

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. 5).

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. 5 and 6).

Ayal, A. (2011). Harmful freedom of choice: Lessons from the cellphone market. *Law and Contemporary Problems 74,* 91–133. (ref. p. 2 and 7).

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. 23).

Busemeyer, J. R. and J. T. Townsend (1993). Decision field theory: A dynamic cognition approach to decision making. *Psychological Review 100*(3), 432–459. (ref. p. 5).

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. 4).

Carlin, B. I. (2009, March). Strategic price complexity in retail financial markets. *Journal of Financial Economics 91*(3), 278–287. (ref. p. 2 and 6).

Chioveanu, I. and J. Zhou (2009). Price competition and consumer confusion. *MPRA Paper No. 17340.* (ref. p. 2 and 6).

DellaVigna, S. and U. Malmendier (2006, June). Paying not to go to the gym. *American Economic Review 96*(3), 694–719. (ref. p. 2 and 6).

Eaton, J. W. (2002). *GNU Octave Manual.* Network Theory Limited. (ref. p. 1 and 16).

Ellison, G. (2005, May). A model of add-on pricing. *The Quarterly Journal of Economics 120*(2), 585–637. (ref. p. 2 and 6).

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. 6).

Ellison, G. and S. F. Ellison (2009). Search, obfuscation, and price elasticities on the Internet. *Econometrica 77*(2), 427–452. (ref. p. 2 and 6).

Freedman, D. H. (2010). *Wrong: Why experts keep failing us.* Little, Brown and Company. (ref. p. 4).

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. 5).

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games.* MIT Press. (ref. p. 5).

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 and 6).

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. 5).

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. 5 and 35).

Gaudeul, A. and R. Sugden (2011). Spurious complexity and common standards in markets for consumer goods. *Economica*. (ref. p. 1, 2, 4, 5, 15, and 22).

Gigerenzer, G. and H. Brighton (2009). Homo Heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science 1*, 107–143. (ref. p. 6).

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. 6).

Gigerenzer, G., P. M. Todd, and The ABC Research Group (2000). *Simple heuristics that make us smart*. Oxford University Press. (ref. p. 7).

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. 23).

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal 7*(3), 388–401. (ref. p. 23).

Huck, S. and B. Wallace (2010). The impact of price frames on consumer decision making. Report 1226, Office of Fair Trading. (ref. p. 6).

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. 6).

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. 6).

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. 6).

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. 3, 8, 15, and 23).

Latty, T. and M. Beekman (2010). Irrational decision-making in an amoeboid organism: Transitivity and context-dependent preferences. *Proceedings of the Royal Society B: Biological Sciences*. (ref. p. 5).

Layard, R. (2005). *Happiness: Lessons from a new science*. Penguin Press. (ref. p. 4).

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

Lusardi, A., O. S. Mitchell, and V. Curto (2010). Financial literacy among the young. *Journal of Consumer Affairs 44*(2), 358–380. (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 and 6).

Miravete, E. J. (2011). Competition and the use of foggy pricing. Working paper, University of Texas at Austin and Centre for Economic Policy Research. (ref. p. 2 and 6).

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. 6).

Payne, J. W., J. R. Bettman, and E. J. Johnson (1993). *The adaptive decision maker*. Cambridge University Press. (ref. p. 7).

Piccione, M. and R. Spiegler (2010). Price competition under limited comparability. Working paper, LSE, Tel Aviv University and UCL. (ref. p. 2 and 6).

Scitovsky, T. (1950). Ignorance as a source of oligopoly power. *The American Economic Review 40*(2), pp. 48–53. (ref. p. 2).

Shestakova, N. (2011). Understanding consumers' choice of pricing schemes. Working paper, University of Vienna. (ref. p. 6).

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics 69*(1), 99–118. (ref. p. 15).

Spiegler, R. (2011). *Bounded rationality and industrial organization*. Oxford University Press. (ref. p. 2, 4, and 6).

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. 5).

Sugden, R. (1989). Spontaneous order. *The Journal of Economic Perspectives 3*(4), 85–97. (ref. p. 6).

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

Thaler, R. H. and C. R. Sunstein (2008). *Nudge*. Yale University Press. (ref. p. 4).

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press. (ref. p. 23).

van Rossum, G. and F. Drake (Eds.) (2001). *Python Reference Manual*. PythonLabs. (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. 6 and 14).

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. 6).

A Index of abbreviations and notations

By order of appearance:

CS: Common standard

IS: Individuated standard

LPCS: Lower priced CS offers

HPCS: Higher priced CS offers

3-menu: A menu composed of three offers

6-menu: A menu composed of six offers

B: Budget given to the consumer to buy paint

up: Unit price

s: Size of the area that can be covered by an offer

p: Price of an offer

Na: Naive, as in naive consumer, that is, a consumer that follows the Naive rule

SF: Signal-First

DE: Dominance editing

Th: Threshold

v: Value of the threshold.

B Instructions

Welcome to this experiment!

1. General rules/proceedings

During the experiment you are not allowed to talk to other participants. Please switch off your mobile phone. If you have any questions, please raise your hand and refer directly to the experimenters. One of the experimenters will then answer your question in private.

Please read these instructions carefully, as your payment will depend on the decisions that you make during this experiment.

On your desk you will find this instruction sheet, a pen, paper, and a receipt. You can take notes at any time; the receipt will only be used for your personal payment at the end of this experiment. During the experiment, we will not speak of Euro but use ECU (*Experimental Currency Units*) as a currency instead.

The amount of ECU you earn during the experiment will be converted into Euro at the end of the experiment using the following conversion rate: **0.8 € = 1 ECU**. For example, if your earnings amount to 12 ECU, you will receive 9.60 €. The final payment will be rounded up to the nearest 10 cents.

All participants will remain anonymous, i.e. after the experiment, no one –neither other participants nor the experimenters – will be able to associate your personal information with your decisions or your earnings.

2. The Experiment

This experiment consists of several tasks. At the beginning of each task, you are endowed with 60 ECU to buy grey paint from a shop in order to paint a specific, given area. Each shop gives a choice between various offers. Each of them is structured in the same way, i.e. it consists of a given shape and its corresponding price. In each offer, the grey shape on display represents the fraction of the total area (which needs to be painted) that you can paint with this specific offer.

Figure 1 presents the three different offers you are given by a shop. **Figure 2** shows the six different offers made by another shop. The total area which you have to paint is represented by the white square surrounding each of the shapes. The light grid is provided to help you with your task.

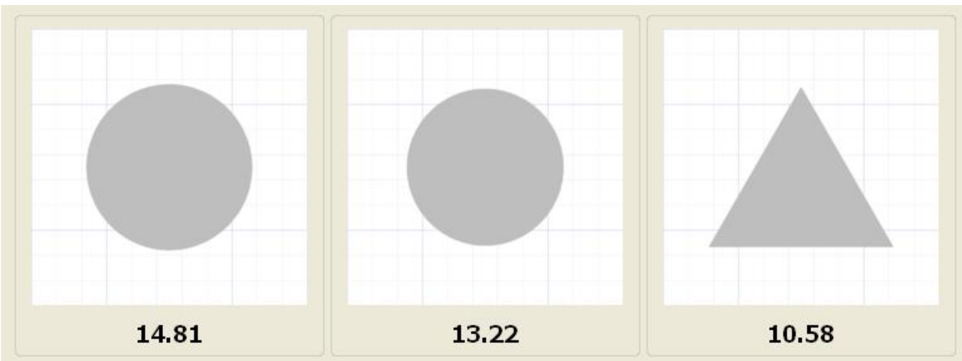


Figure 1

Once you have chosen one of the offers and submitted your choice, the computer will calculate how much paint you need to cover the entire area (the white square) and will also buy the colour for you. The amount of your initial endowment that you do not spend for buying the paint is yours to keep.

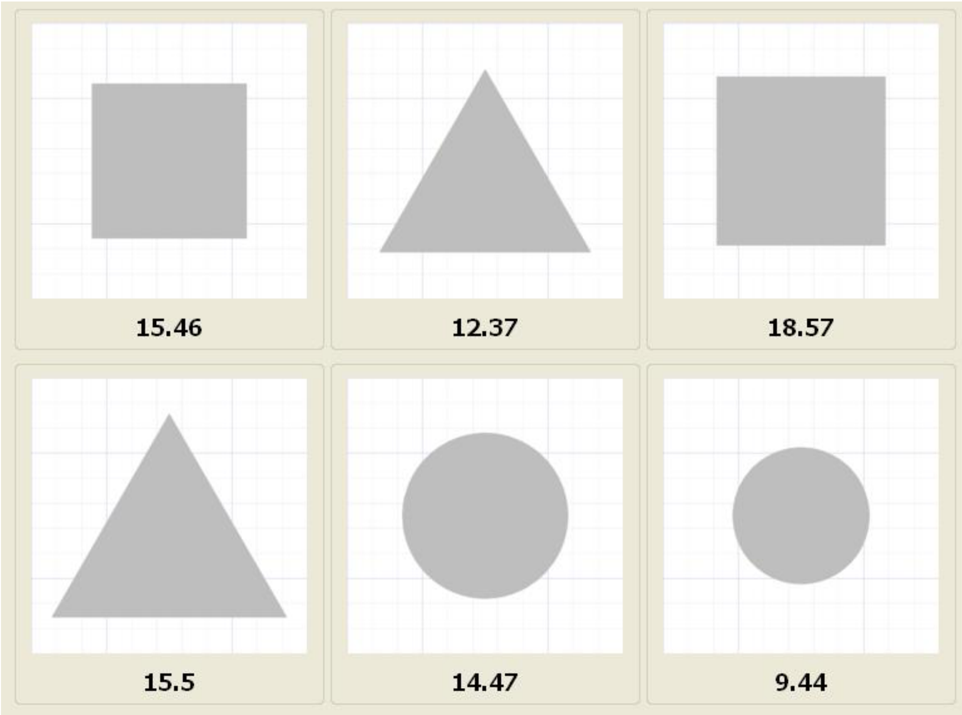


Figure 2

3. Examples

The following examples should help you understand how the computations made by the computer work in detail. Suppose you are confronted with the offers in **Figure 3** and the total area you are supposed to paint is 100m².

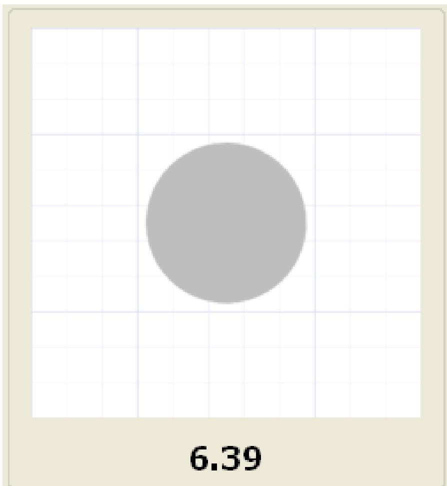


Figure 3

In order to paint the area covered by the grey circle, you pay **6.39 ECU**.

However, this circle only covers an area of **13m²**. As you need to paint a square which is 100m² in size, the computer calculates how much paint you actually need for this offer.

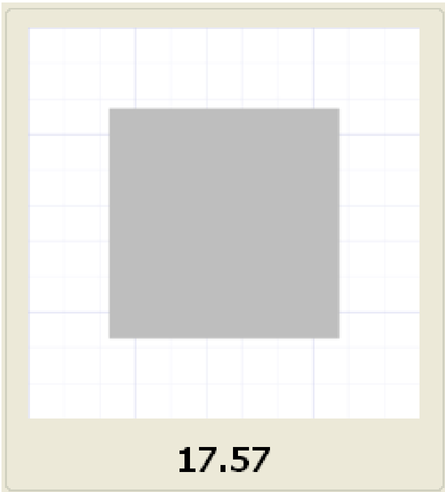
In this case, this amounts to **100/13 = 7.7** paint buckets.

Hence, the total price you have to pay for painting the white square amounts to:

6.39 x 7.7 = **49.2 ECU**

Keeping in mind your initial endowment of **60 ECU**, your earnings result as follows:

60 – 49.2 = **10.8 ECU**.



In order to paint the area covered by the grey square, you pay **17.57 ECU**.
However, this square only covers an area of **34m²**. As you need to paint a square which is 100m² in size, the computer calculates how much paint you actually need for this offer.

In this case, this amounts to **100/34 = 2.94** paint buckets.

Hence, the total price you have to pay for painting the white square amounts to:

17.57 x 2.94 = **51.7 ECU**

Keeping in mind your initial endowment of **60 ECU**, your earnings result as follows:

60 – 51.7 =**8.3 ECU**
A separate pop-up dialog will automatically appear and will tell you the results of each task (see **Figure 4**) including your possible earning of this task; clicking ‚OK‘ will start the next task.

Figure 4

You have at most two minutes for each task and can only submit a choice at least ten seconds after you started it. In case you made a choice after two minutes (i.e. clicked on one of the offers), but failed to submit the offer in time (by clicking 'OK'), the computer will nevertheless treat your selected offer as if you had submitted it. In case you did not make any choice after two minutes, you will be paid 3 ECU for this task (if this task is chosen as relevant for you payment).

You will be faced with 36 different tasks with 3 offers, and 44 with 6 offers. At the end of the experiment, only one of the 80 tasks will be randomly selected and you will be paid according to your earnings in this specific task.

4. Questionnaire and Quiz

Once you completed the 80 tasks, you will be asked to answer a few questions:

1. Please fill in a simple questionnaire. The answers you submit will be treated confidentially and no data will be disclosed.

2. Please compare different shapes with each other. You have one minute for each of the four comparisons.

3. Please perform some computations. There will be 3 sets of computations and you will have one minute for each.

4. Please solve a number of problems. There will be 4 problems, and you have 2 minutes for each.

After you completed all the tasks, please raise your hand to signal the experimenters that you finished the experiment and we can start with your payment. One of the experiments will then come to your cabin and ask you to draw a chip out of a bag with 80 chips (which are numbered 1 to 80). This chip will correspond to the task that you will be paid for. The experiment will then enter the number of the chip on your screen and the computer will automatically tell you, how much you earned in this task. Please fill in this amount as well as your name and signature the receipt that you find on your desk. Afterwards, please raise your hand to signal the experimenters that you are finished filling out your receipt. After you received your payment, the experiment is finished and you can leave the laboratory.

Thank you very much for participating in this experiment!

C Screenshots

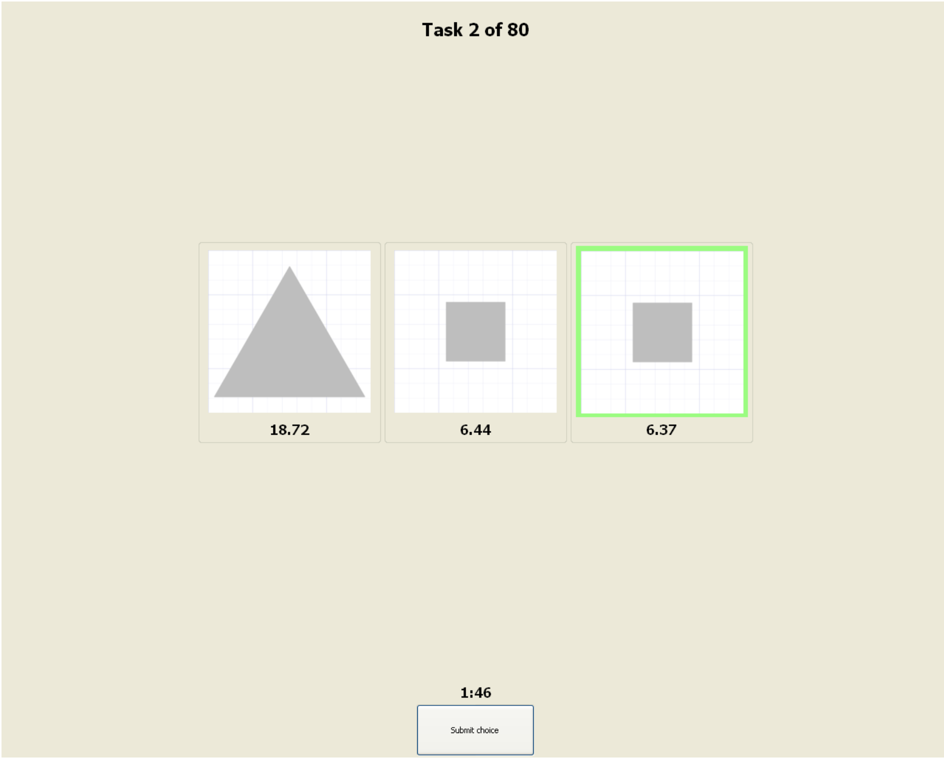


Figure 9: Screenshot, 3-menu with a common standard and a selected offer

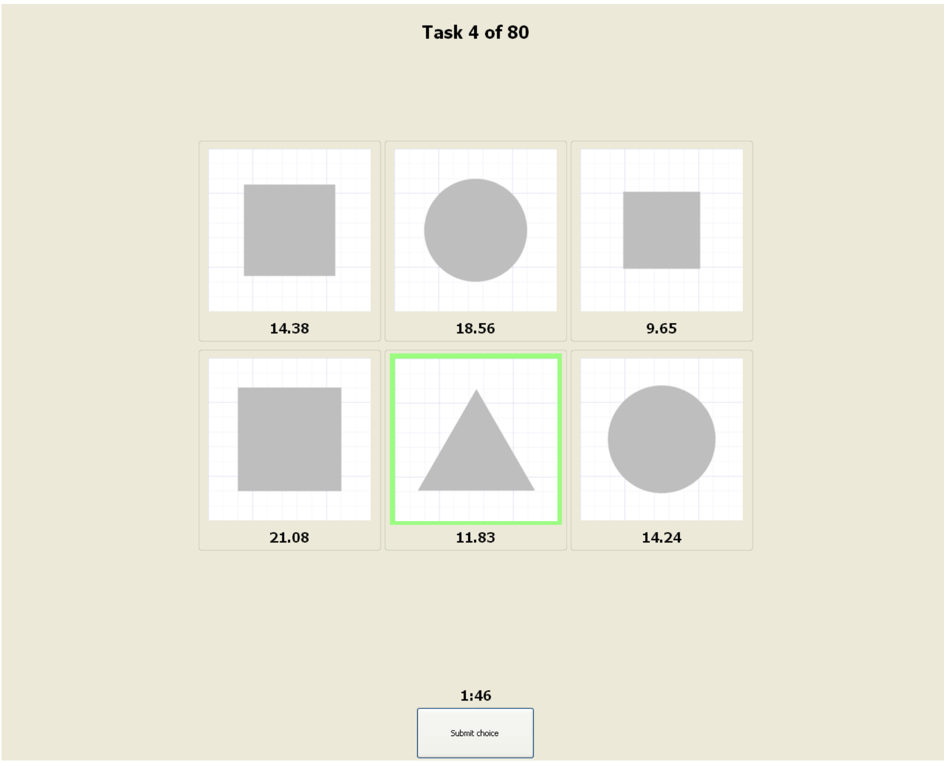
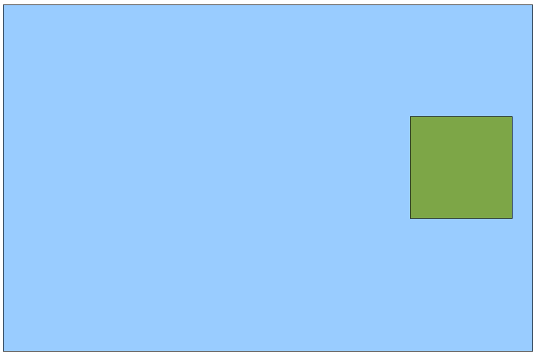


Figure 10: Screenshot, 6-menu without a common standard and a selected offer

D Control tasks

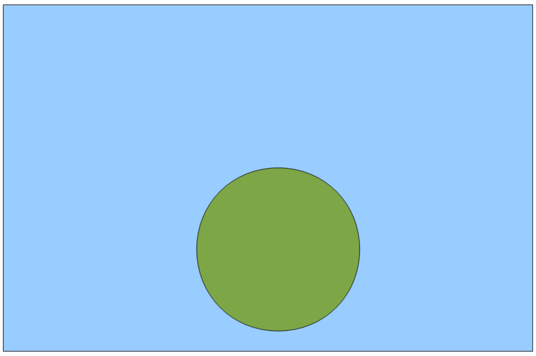
D.1 Shape comparison task

- Question 1 (one minute)



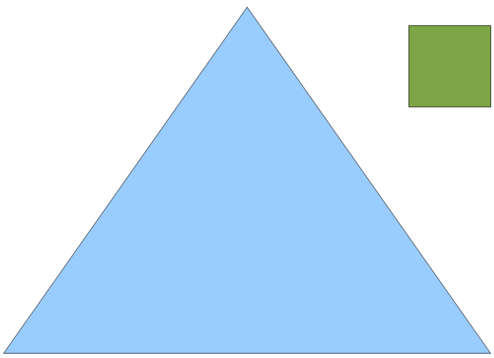
How many times bigger is the area covered by the rectangle compared to the area covered by the square?

- Question 2 (one minute)



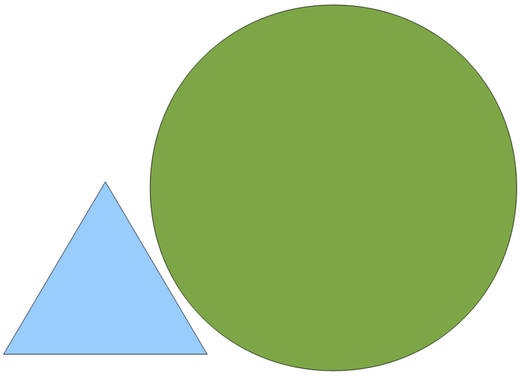
How many times bigger is the area covered by the rectangle compared to the area covered by the circle?

- Question 3 (one minute)



How many times bigger is the area covered by the triangle compared to the area covered by the square?

- Question 4 (one minute)



How many times bigger is the area covered by the circle compared to the area covered by the triangle?

D.2 Simple computations

Time given: 1 minute, allow for not answering some questions, put all questions at same time to be answered in box next to each problem

- Question 1 (one minute)

88 – 45; 10 + 30; 57 – 43; 9 x 6; 3 x 7; 8 + 45; 65 – 11; 2 x 5; 8 + 12.

- Question 2 (one minute)

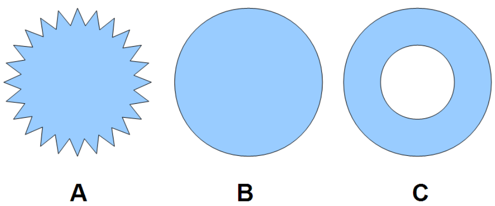
276 + 177; 12 / 4; 106 – 85; 18 / 6; 4 x 10; 188 – 64; 106 + 122; 8 x 7.

- Question 3 (one minute)

70/10; 892-179; 8*8; 363+93; 77/11; 9*5; 642-193; 265+108.

D.3 Problems

- Problem 1 (two minutes)



Which of the figures has the largest colored area?

- Problem 2 (two minutes)

A pizzeria serves two round pizzas of the same thickness in different sizes. The smaller one has a diameter of 30 cm and costs 3 euros. The larger one has a diameter of 40 cm and costs 4 euros.

Which pizza is better value for money?

1. The smaller one
2. The larger one
3. Both are the same value for money

- Problem 3 (two minutes)

Nick wants to pave the rectangular patio of his new house. The patio has length 5 metres and width 3 metres. He needs 80 bricks per square metre. How many bricks does Nick needs for the whole patio?

- Problem 4 (two minutes)

You can buy $1.40 with one euro. How many dollars can you buy with 50 euros?