How does the frame of menu influence choices? -Investigating Preference (In)Consistency in Two-Stage Decision Problems

Lihui Lin[1] and John Hey[2]

Abstract

This paper experimentally investigates several theories of decision-making in a two-stage context - one in which the decision-maker (henceforth DM) must first choose a menu from a set of menus, and secondly must choose an item from the chosen menu. While standard decision theory analyses this problem using backward induction (and assuming stable preferences), several new theories postulate that the DM might anticipate that his/her preferences may change, and that the DM might well take this into account when deciding which menu to choose. Leading amongst such new theories is self-control theory, which incorporates the notion that the DM might anticipate temptation at the second stage, and hence might exercise self-control at the first stage to avoid being tempted. The other theories suggest that the DM might prefer flexibility at the second stage, and this may affect the choice at the first stage. These can all be observed in life. This suggests that the composition of choices in menu, and *frames* of menu choices presented to the DM, may affect choice of menu. Our experimental show that placing choices in different meus and the frames of available menu choices lead decision rule changes under this two-stage context.

Key words: two-stage decision-making; frame effect; preference inconsistency; preference uncertainty; temptation; experiment.

AEA codes: C91, D81,

Acknowledgements: We are grateful to the Leverhulme Trust for generously funding this research with Leverhulme Emeritus Grant EM-2019-026/7

[1] ll1654@york.ac.uk, Department of Economics, University of York, York, YO10 5DD, UK.

[2] john.hey@york.ac.uk, Department of Economics, University of York, York, YO10 5DD, UK.

1. Introduction

Consider an everyday decision problem: in the morning, the decision-maker (DM) has to choose a restaurant (from a set of restaurants) to go to in the evening; in the evening, the DM arrives at the chosen restaurant and has to choose an item from the menu.

Standard economic theory provides a clear solution to the problem (assuming that the DM has well-defined preferences over all menu items in all restaurants under consideration and that the quality of the cooking is the same in all restaurants): the DM should work backwards. First, the DM should imagine herself[3], for each restaurant under consideration, arriving at that restaurant in the evening, and choosing the best item from that restaurant. Then (in the morning) she should choose the restaurant where the best item is the most preferred. All that matters is the best menu item at each restaurant; any other items (and the numbers and types of them) are irrelevant.

One thing that follows that if one restaurant has a larger menu than another has (and contains the same menu items as the other), then former will be chosen in the morning. However, this 'solution' appears to be unrealistic: for example, a would-be vegetarian might want to avoid restaurants that serve both meat and fish, for she might fear that she would be tempted in the evening by a meat item (which she is trying to avoid).

A number of recent theory papers have tried to construct models that are more realistic. Common to many of these new theories is the idea that the DM does not have unique well-defined preferences over all menu items in all restaurants under consideration, and instead may have preferences different from her preferences in the morning when she arrives at the restaurant in the evening. These latter are called by some theorists 'temptation preferences'. If this is the case, then the thing of interest to theorists is how the conflict between her '*ex ante*' (morning) preferences and her temptation preferences is resolved.

The literature in this new field is commonly referred to as 'the literature on self-control problems. Most of this literature is axiomatic. A leader in this field is Gul and Pesendorfer (2001), who model the DM as anticipating her dilemma in the evening and taking into account the cost to herself of exercising self-control. This may lead to a desire for commitment at the choice of menu stage. An earlier contribution

[3] For 'herself' read 'herself' or 'himself', and similarly, *mutatis mutandis*; for 'she' and 'her'.

by Kreps (1998), suggests that uncertainty about future preferences may instead lead to a desire for flexibility. Both behaviours can be observed in the same context. The existence of tempting choices will affect the individual's desire for current versus future consumption (that is, time inconsistency) or the uncertainty about future preferences, which we term as a taste shock. Time-inconsistent preferences generate demand for commitment, but uncertainty about future preferences generates demand for flexibility.

This paper applies the above two lines of literature to experimentally study how menu frames influence consumer behaviour. In literature, framing refers to the phenomenon of "simple and unspectacular changes" in the presentation of decision problems, leading to changes in choice (Kühberger 1998). We focus on the framing effect of menu choices with the application of temptations. We design an experiment to investigate how menu frames influence subjects' decision rules by identifying which of the various stories (the standard model, self-control and flexibility) best describe subjects' decision rules. We assume that subjects differ in their preferences, so we first infer subjects' ex ante preferences. We then observe the choices of our subjects in this two-stage decision problem (from which we can infer whether they suffer from temptation and how they resolve the temptation). Menusets are carefully designed so we can identify the demand for flexibility and commitment. Lotteries are used as the final objects of choice, allowing us to measure quantitatively temptation and the decision rules of the subjects.

This paper is organised as follows. Section 2 discusses the main literature in this field. Section 3 describes the framework and gives motivating examples for our research. Section 4 describes the experimental design and section 5 the econometric specifications. Section 6 discusses and interprets the experimental results. Section 7 concludes with a summary of the results, a discussion of findings and implications.

2. Related literature

Our analysis relates primarily to a class of preferences over menus that addresses the role of temptation within a menu in the presence of self-control and flexibility.

An early paper on the role of temptation in menu choices is that of Gul and Pesendorfer (2001). This modeled a decision maker who anticipates being tempted and executes self-control to resist the temptation at the choice of menu stage, thus incorporating the disutility of temptation. There are other models following this line but adopting different assumptions. A generalization of GP is Chatterjee and Krishna (2000) - henceforth CK. While GP incorporates a cost of self-control when the DM deliberately excludes possible temptations, CK incorporates a cost from the risk of succumbing, that is, from 'random indulgence' rather than costly self-control: the DM implements a dual-self-evaluation which are determined by the long-term normative preference and the temptation-driven preference incorporating the fact that the individual considers the possibilities of both selves. Most models in this line derive the idea of a commitment demand that allows the DM to exclude temptations from the menu.

Empirical evidence shows that excluding *ex-post* opportunities is not always desired; a preference for flexibility is also widely observed. A preference for flexibility corresponds to strictly preferring a restaurant that serves two options to a restaurant that serves only one or the other. This preference for flexibility was first modelled by Kreps (1979). Kreps modeled this preference by considering the agent's choices over menus of options by anticipating the possibility to choose from the chosen menu, where the chosen menu will be her choice set at a future date. Dekel, Lipman, and Rustichini (2001; henceforth DLR) enriched Kreps's domain of choice from menus of deterministic alternatives to menus of lotteries. The key feature of self-control (as distinct from flexibility) is the desire to keep commitment and eliminate possible temptations.

These two lines of research have been applied to optimal contract design. Even though they suggest opposite behaviours, both of them are observed in real life: hence, the value of commitment and of flexibility should both be considered. Manuel and Iván (2006) state that the DM can act upon the taste shock driven by temptation and trade-off between commitment and flexibility. Philip and Gustav (2013) analyze a contract on a consumption-saving problem considering demand for commitment and demand for flexibility. But few investigate how the menuset design influences the choice rules regarding the demand for commitment and the demand for flexibility, nor answer how sellers should respond to this demand heterogeneity (though this latter will not be our concern).

Our research context is close to that of Esteban and Miyagawa (2022) who examine the optimal design of menusets when investigating the pricing strategy of sellers, and how they can benefit from offering multiple menus and adding temptation into menus. However, we do not try to find the optimal design of menusets; instead, we are interested in how different menusets influence consumer behaviour. This has a similar spirit to Yuval's research on 'contracts with frames'. Yuval (2018) justified product menusets with frames as temporarily increasing the attractiveness of some products and derives an optimal menuset design leading to higher profits. However, their context focusses on a choice *of* a menu situation. Ours explores the frames of menus and the final choice *from* the chosen menu.

Menus can be interpreted either literally or as an action concerning an opportunity set, such as signing up to a contract or choosing a service package (product bundles that will affect subsequent opportunities). Our research mainly makes a contribution to product assortment strategies incorporating consumer heterogeneity. There is growing evidence that consumers tend to focus on the set of options they happen to observe in a particular context (for example, the items on the shelf) and use that set to determine which, if any, of the options is attractive. This tendency can have significant implications for retailers, since consumers typically consider only a subset of the entire product assortment. Accordingly, the configuration of subsets can be the key determinant of purchase decisions.

3. Framework

We consider an environment where sellers produce a variety of goods. They produce a collection of goods $\{e_1, e_2, \ldots, e_I\}$ which are differently preferred by different customers. They also produce some goods that are less desirable to most customers $\{t_1, t_2, \ldots, t_J\}$ but could be considered desirable under some circumstances. Let us term these as *temptation goods*.

In our experiment, we need to choose some particular goods for our subjects to choose. Moreover, we need to be able to measure the attractiveness of these good to our subjects. We therefore chose as our 'goods' *lotteries*. The relative attractiveness of the goods is given by their riskiness; if we know the risk-attitude of our subjects, we can rank the goods/lotteries by their risk-attitude. We assume that our subjects differ in terms of their risk-attitude: there is preference heterogeneity amongst our subjects.

In response to this heterogeneity, sellers often use product assortment[4] strategies to offer consumers different menus or service packages. Buyers need to choose a particular choice set first and make a decision from this choice set.

Consider two motiving examples.

Case 1:

Consider a frugal traveler who is planning a trip and needs to choose a hotel room. During her trip, she expects to use the room only to sleep. Hence, she prefers a standard level room $C = \{e_1, e_2\}$. So, she only checks this category and books one and wants to avoid the luxury type $T = \{t_1, t_2\}$, as she is worried if she will be tempted by luxury fancy rooms. The hotels want to increase the chance of selling a luxury room, so they offer another flexible service package allowing travelers to delay commitment, that is a menu $F = \{e_1, e_2, t_1, t_2\}$. The hotel hopes that travelers can be tempted to choose one of t_1 or t_2, and therefore offers her the menu $F = \{e_1, e_2, t_1, t_2\}$ to enable her possibly to be attracted by the luxury types after arriving the hotel.

Case 2:

Consider a big vehicle company owning three stores and which wants to promote a new artificial intelligence automobile. It is a new technology with lots of fancy functions. But it is an early stage for this kind of technology, and consumers may feel it is too risky to buy them. If the company's strategy is to arrange the products in three stores: $C = \{e_1, e_2\}$, $T = \{t_1, t_2\}$, and $F = \{e_1, e_2, t_1, t_2\}$, consumers may not go into $T = \{t_1, t_2\}$ so that this store will not be profitable. But if the company places the new products in every store: $C = \{e_1, t_1\}$, $T = \{e_2, t_2\}$ and $F = \{e_1, e_2, t_1, t_2\}$, consumers will get information about the new model whichever store they visit.

[4] Product assortment, sometimes referred to as merchandise mix, refers to the variety of products that a retailer stocks and sells.

We ask: does buyers' behaviour differ in the two cases? If it does, how do they make choices? To answer these questions, we should understand how buyers make choices in these cases.

Buyers: in our context, we assume that the buyers have preferences over lotteries, and that their preferences are uniquely described by their *risk attitude*[5] *r*. We need to distinguish between their *ex ante* (or long-term, or normative) preferences with parameter r_u, and their temptation (or *ex post* or *ex ante* undesirable) preferences with parameter r_v. It is the fact that these may be different that causes the possible conflicts in their decision-making. We define the preferences *outside* the menu-choosing context r_u, as their *ex ante* normative preferences. These can be understood as the buyer's long-term goal before making choices in a menu context. However, when facing choice in a menu-choosing context, a taste shock may be induced by the existence of some tempting choices. We assume that tempting choices are *ex ante* unattractive, but they are attractive under r_v. For instance, a DM says *ex ante,* that she is not a fan of beef steak and prefers to avoid it: but, in the restaurant, she is tempted to choose steak by the smell in the restaurant. We understand the menusets themselves create an environment in which the individual's desires conflicts with their initial desires. When they make a choice of menu from the available menus, they apply different decision rules and a different menu utility function to evaluate menus.

Different types of buyers have different rules; let us use q to denote the type of a buyer (we give specific cases below). We introduce some notation: let $U(r_u, r_v, q)$ denote the utility function of a buyer of type q, in the choice of menu stage with *ex ante* preference r_u and ex-post preference r_v. We identify three diferent types. We label these as those who follow standard theory (ST), those who care about flexibility (PF), and those who follow self-control theory (SC). We get the following three decision rules over menus:

[5] Later we will assume that their utility function is of the CRRA form with parameter *r*.

Standard theory (ST): The **DM** applies standard backward induction to solve the two-stage problem: first, decide what is optimal to do at the second stage (given what has been decided at the first stage), and then, in the light, of that decide what is optimal to do at the first stage.

Preference for flexibility (PF): The **DM** anticipates the probabilities of each choice she will make in future and how much utility these choices will provide. This implies that the menu with more possible choices under DM's consideration, may be more preferred.

Self-control (SC): The **DM** will anticipate a preference conflict driven by temptation and will tend to avoid the temptations in the menu. It implies that a smaller menu size without temptation may be preferred, as the existence of temptation in the menu will decrease the utility of the menu.

Menusets

We need some preliminaries. Let us denote the lotteries to be inserted into the menus in the menusets by l_i ($i=1,\dots$,I). The *ex ante* preferences will determine a ranking of these; we suppose that this is a complete ranking and we denote the lotteries ranked in this way by e_i ($i=1,\dots$,I). (It follows that, *ex ante*, $e_1 \succ e_2 \succ ...e_I$, where $\succ$ indicates strict preference. Similarly, the *ex-post* preferences will determine a ranking of the lotteries; let us suppose again this is a complete ranking and denote the lotteries ranked in this way by t_i ($i=1,\dots$,I). (It follows that *ex-post* $t_1 \succ t_2 \succ ...t_I$)

Now, let us talk about *menusets*. We define a menuset as a collection of menus. We restrict ourselves here to menusets of size three, so that there are three menus in each menuset. A menu contains 2, 3 or 4 lotteries. The composition of a menuset depends upon the *frame*.

Let us denote a menuset of frame *f* by *S(f)*. The frames *f* differ in terms of their flexibility and the location of the *ex ante* undesirable choices in different menus. We define a menu containing more choices as one with a higher degree of flexibility. Motived by two typical cases, we consider four possible menusets (the rationale for which we will explain later):

Two of them are designed following the two above-mentioned motivating exmples, that is, $S(f_1)=[\{e_1,e_2\},\{t_1,t_2\},\{e_1,e_2,t_1,t_2\}]$ and $S(f_2)=[\{e_1,t_1\},\{e_2,t_2\},\{e_1,e_2,t_1,t_2\}]$[6]. The others consider higher flexibility cases, with more *ex ante* undesirable choices.

$S(f_3)=[\{e_1,e_2\},\{t_1,t_2,t_3\},\{e_1,e_2,t_1,t_2\}]$ and $S(f_4)=[\{e_1,t_1\},\{e_2,t_2,t_3\},\{e_1,e_2,t_1,t_2\}]$.

Our notation implies that e_1 and e_2 are the two most attractive lotteries according to *the ex ante* preferences r_u; and that t_1, t_2 and t_3 are three most attractive lotteries according to the *ex post* preferences r_v. t_1 , t_2 and t_3 are different source of *temptation* to the DM.

The choice of the menu from the menuset is determined by the *type* of DM. We can use the types to produce the Table 1. If there is a single entry, then that is the optimal choice; if there are two or more entries the DM is indifferent between them and can be presumed to choose at random. If the Table says 'depends' then the optimal decision depends upon the magnitudes of the risk-aversion parameters, and may vary from individual to individual.

Frame	f_1	f_2	f_3	f_4
Sets of menus	$[\{e_1,e_2\},\{t_1,t_2\}, \{e_1,e_3,t_1,t_2\}]$	$[\{e_1,t_1\},\{e_2,t_2\}$ $\{e_1,e_2,t_1,t_2\}]$	$[\{e_1,e_2\},\{t_1,t_2,t_3\}$ $\{e_1,e_2,t_1,t_2\}]$	$[\{e_1,t_1\},\{e_2, t_2,t_3\},\{e_1,e_2,t_1,t_2\}]$
ST	$\{e_1,e_2\}$ $\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$ $\{e_1,e_2,t_1,t_2\}$	$\{e_1,e_2\}$ $\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$ $\{e_1,e_2,t_1,t_2\}$
PF	$\{e_1,e_2,t_1,t_2\}$	$\{e_1,e_2,t_1,t_2\}$	$\{t_1,t_2,t_3\}$ $\{e_1,e_2,t_1,t_2\}$	$\{t_1,t_2,t_3\}$ $\{e_1,e_2,t_1,t_2\}$
SC	$\{e_1,e_2\}$	Depends on risk-aversion parameters	$\{e_1,e_2\}$	Depends on risk-aversion parameters

Table 1: The different frames and the menu choice of each type

Let us start by explaining the first column, f_1: given that an ST type uses backward induction, she starts by deciding the best choice from each menu: these are e_1 from $\{e_1,e_2\}$,either t_1 or t_2 from $\{t_1,t_2\}$, and e_1 from $\{e_1,e_2,t_1,t_2\}$. However, as e_1 is preferred to either or both of t_1 and t_2, it is best to choose either the

[6] Some may wonder why not design the menus following more common examples in most papers like $\{e_1\},\{t_1\},\{e_1,t_2\}$. The reason we design in our way is so that subjects were given a richer and more realistic choice set, and so that we can distinguish between the different decision rules.

first or third menu. The PF type prefers $\{e_1,e_2,t_1,t_2\}$ because it contains two other menus. The SC type chooses $\{e_1,e_2\}$ to avoid the utility loss from being tempted.

As regards the second column, the decisions of the ST and PF types follows with a similar argument to the first column. However, the decision for an SC type is not trivial. The menuset under f_2 does not offer any menu choices that exclude *ex ante* inferior options (that is, the commitment menu $\{e_1,e_2\}$), thus the disutility of resisting temptation needs to be included in every menu evaluation. A SC DM needs to trade off the utility of the *ex ante* most preferred option in the menu and the disutility of resisting temptation. In the other words, menu utility and hence the menu choice depends upon the difference between the utility of the *ex ante* most preferred option in the menu and the disutility of resisting temptation.

Frames f_3 and f_4 are analogous to frames f_1 and f_2 respectively, with only one difference: that one menu has more temptations. Adding *ex ante* inferior options into menus will not influence their menu utility for the ST and SC decision maker. However, as flexibility is increased (by adding new options), the PF decision maker may behave differently. Either $\{t_1,t_2,t_3\}$ or $\{e_1,e_2,t_1,t_2\}$ may be preferred conditional on the anticipated possibility distribution to choose these options and the corresponding utility of each option.

Overall, the fundamental difference between the three different types is the DM's attitude towards the role of *ex ante* undesirable choices and the possible conflict between r_u and r_v. Both PF and SC types perceive a preference conflict triggered by the presence of *ex ante* undesirable choices (but choices preferred by r_v). The PF *type* positively perceives them as flexibility for future opportunities; the SC typ*e* negatively responds to them by excluding the future possibilities, while the SD typ*e* does not perceive preference conflicts.

Choice from menu – $S(f_1)$ satisfies the flexibility and commiment demand for PF and SC decision makers respectively. $S(f_2)$ mixes temptation choices with normatively preferred choices and makes the commitment menu unavailable. In this context, the DM has entered an environment where tempting choices exist. Observing the choice of menu under these frames from different types of DM and tracing the corresponding final choice from the menu can answer the main question of this project – how menu frames influence consumer behaviour.

4. Experiment design

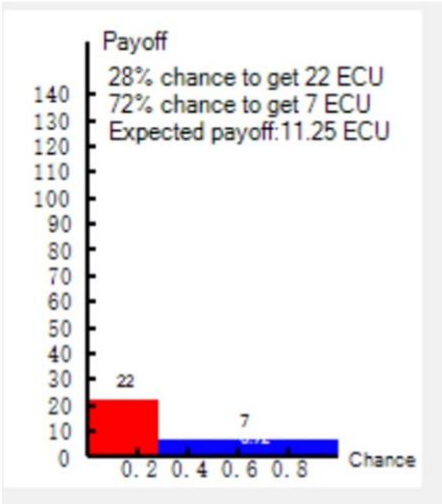
4.1 Experimental Procedure

Our experiment consisted of two parts. In the first Part, we identified the subjects' *ex ante* preferences using (our slight modification of) Holt-and-Laury's price list method and hence estimated the subjects' *ex ante* risk attitude. (We assumed throughout that all preferences were CRRA Neumann-Morgenstern.) In the second Part, subjects were asked to choose one menu from a set of menus, and then choose an option from the chosen menu. Importantly, we gave different subjects different sets of menus and different menus depending on their *ex ante* risk preferences; this embodies the idea of a personalised offering based on an *ex ante* market survey (as current online shopping websites do).

Part 1 consisted of 28 tasks, and Part 2 consisted of 60 tasks. As we shall discuss, the menus and menusets offered to subjects in part 2 varied across subjects according to their evaluation of the singletons in part 1. At the end of the experiment, one of the total of the 88 tasks was chosen at random, and the subject's decision on that task was 'played out'.

As we have already noted, we used lotteries as the menu items. All the lotteries in the experiment were simple ones, each with just two outcomes. We used the representation of Figure 1.

Figure 1: the representation of lotteries



4.2 Temptation

We need to implement possible temptations in our menusets. We employ two different types of temptations: risk-free lotteries and very-risky lotteries. Both may create a conflict between the *ex ante* preferences[7] and menu choices in Part 2. We refer to very risky lotteries as gamble temptations, as they may stimulate an urge to gamble.

Both risk-free and very-risky menu items have a lower expected payoff than the other menu items. Risk-free items are a certainty of £8, £9 or £10. The very-risky items' high payoffs are the triple of the high-payoffs for the other items, and they have a low probability (less than 0.05); while the low payoffs are lower than any normal items' low payoffs.

We designed the menusets in a particular way so that the *ex ante* preferred choices and *ex ante* unpreferred choices can be clearly determined. Risk is the single measurement of the attractiveness of lotteries. We generated 30 sets of lotteries (each set has 7 lotteries) for 60 menu choice tasks according to particular riskiness. It follows that each set of 7 lotteries can be ranked according to particular value of risk preference r. We gave each subject all 30 menusets twice – giving a total of 60 menusets in Part 2.

4.3 An illustration of the lotteries included in different menusets

[7] In Part 1 none of our subjects were completely risk-averse or completely risk-loving.

We start with a 7 by 210 matrix, each cell referring to a lottery. The attractiveness of lotteries *in each row* is ranked according to each subjects' r_u. There are 7 entries/lotteries in each row; each row defines the lotteries to be put into a menuset (the actual construction of the menusets is described later). The rows are constructed in blocks of 10. The first block of 10 rows is all of the same type, though the actual lotteries differ (but the five *e*'s are still defined as the five most preferred according to the *ex ante* preferences) and the *t*'s are all *very risky* lotteries (but differ from row to row). Similarly, the second block of 10 rows are all of the same type, though the actual lotteries differ (but the five *e*'s are still defined as the five most preferred according to the *ex ante* preferences) and the *t*'s are all *risk-free* lotteries (but differ from row to row). Likewise, the third block of 10 rows are all of the same type, though the actual lotteries differ (but the four *e*'s are still defined as the four most preferred according to the *ex ante* preferences) and one of the *t*'s is a *risk-free* lottery and the other two are *very-risky l*otteries (but differ from row to row).

$$\begin{pmatrix} & e_{1,1} & e_{1,2} & e_{1,3} & e_{1,4} & e_{1,5} & t_{1,1} & t_{1,2} \\ & & & & \text{¼¼} & & & \\ e_{11,1} & e_{11,2} & e_{11,3} & e_{11,4} & e_{11,5} & t_{11,2} & t_{11,3} & \\ & & & \text{¼¼} & & & & \\ e_{21,1} & e_{21,2} & e_{21,3} & e_{21,4} & t_{21,1} & t_{21,2} & t_{21,3} & \\ & & & \text{¼¼} & & & & \end{pmatrix}$$

We note that though there are repetitions within each block since the actual lotteries differ; this was to stop subjects memorising the lotteries.

We summarise the construction of the menusets:

1. Menusets with very-risky temptations
2. Menusets with risk-free temptations
3. Menusets with multiple temptations

Now we discuss the different *frames.*

The implied menusets (depending upon the *Frame*) are

Frame 1: $S(f_1) = [\{e_1, e_2\}, \{t_1, t_2\}, \{e_1, e_3, t_1, t_2\}]$

Frame 2: $S(f_2) = [\{e_1, t_1\}, \{e_2, t_2\}, \{e_1, e_2, t_1, t_2\}]$

In this, e_1 and e_2 are the most preferred according to Part 1 preferences, and t_1 and t_2 are two risk-free items in a menuset with risk-free temptations, and are very-risky items in a menuset with very-risky temptations. The difference between the two Frames is cosmetic according to Standard Theory, but may have an effect with one of the non-standard theories. These are both repeated ten times with different lotteries.

Now we need to include menusets with multiple temptations. This leads to two more frames:

Frame 3: $S(f_3) = [\{e_1, e_2\}, \{t_1, t_2, t_3\}, \{e_1, e_3, t_1, t_2\}]$

Frame 4: $S(f_4) = [\{e_1, t_1\}, \{e_2, t_2, t_3\}, \{e_1, e_2, t_1, t_2\}]$

In these, t_1, t_2 and t_3 are one of the certainties and two of the riskiest lotteries. As these are possible temptations, we denote these with a '*t*'.

5. Econometric specification

5.1 Menu utility function

We used three different menu utility functionals to identify different types of DM in our experiment. The simplest, of course, is that with Standard Theory: an individual with a standard functional simply uses their normative preference to evaluate all available options among the three menus, then the utility of menu is given by the maximised utility of lotteries in the menu, which is the one she plans to choose from the menu.

Standard theory:

$U_{SD}(A) = \max_{x \in A} u(x)$ where x is an option in menu A. (1)

Recall that we assume CRRA utility. Thus, utility is determined by the single parameter of risk preference. We assume that standard theory subject has consistent preferences across stages and the inconsistency in behaviour is interpreted as noise. So, the *ex ante* preference r_u which is inferred from part 1 will be applied to measure the menu utility.

As for the self-control, there are rich research in this vein. They offer different solutions to model the preference over menus. One of the most popular is GP's costly self-control model. Their axioms yield a representation that identifies the conflict between *ex ante preference u* and anticipated ex post *temptation preference v* and incorporates a cost of self-control. To some extent, our experiment context is more general without addressing the cost of self-control as in most self-control experiments. For example, Toussaert (2018) who designed an experiment to test GP, implements the cost of self-control by offering additional earnings to read a sensational story during a tedious attention task for which subjects received payment. In our experiment, the self-control and demand of commitment is triggered by preference inconsistency driven by designed temptation where the DM anticipates the risk to succumb to temptation. Alternatively, another model in terms of self-control emphasizing on the anticipated risk to departure from normative preference is the dual self with a stochastic story proposed by Chatterjee and Krishna (2000)(CK). In contrast to using strong self-control with cost involved, DMs are modelled as evaluate the possibility of being tempted in the future. This model interprets temptation as a systemic mistake in which the second-stage choice could be interpreted as being made by an "alter ego" who appears randomly; DMs at the first stage take into account the probability of being tempted at the second stage.

Self-control functionals

$$U_{SC}(A) = (1-q)\max_{x\in A} u(x) + q\max_{y\in B_v(A)} u(y) \text{ where } B_v(A) \text{ is the set of } v \text{ maximisers in } A$$

(2)

Here q is the probability of being tempted. CK refer to their model as a *dual-self model*, where *u* is the utility function of the long run self and *v* is the utility function of the other self. When making the choice *of* menu, the DM believes she has *q* probability that the other self will dominate at the choice *from* the menu stage; that is, giving into the temptation. Following the spirit of self-control, the representation could be interpreted as an internal battle for self-control with the alter ego where *q* is the probability to lose self-control.

The parameters to estimate in this functional are the anticipated probability q and the preference r_v.

A class of preference over menu representations, which mainly originates from Kreps's preference for flexibility, says that expanding menus by adding alternatives is always desirable. He models a DM who is uncertain about her future preferences and anticipates the probabilities of each choice she will make in future and how much utility these choices will provide.

Preference for Flexibility functionals

$$U_{PF}(A) = \sum q(s) \max u_s(x) \text{where } x \in A \qquad (3)$$

Where $s \in S$ is the subjective state and the $q's$ are the probabilities of anticipated future preferences. Specifically, the utility of a menu *A* is equal to the expected utility of the best option in *A*, with expectations taken over the different possible utility functions indexed by the state *S*, which we refer to as the belief about the preferences at the second stage.

Here the parameters to be estimated are the anticipated probability distribution across states. The estimation difficulties are that the subjective state and the belief on probability cannot be directly observed. Following the spirit of CK's dual self and given the design of our experiment which offers two types of options, we assume there are two states, s_u the normative preference state where has a maximiser determined by r_u and s_v the temptation indulgence state where has a maximiser determined by r_v (that is, the taste shock driven by temptation). Without losing generality, this binary-states assumption does not change the nature of preference for flexibility. For example, a DM makes choice from the menusets $S(f_1) = [\{e_1, e_2\}, \{t_1, t_2\}, \{e_1, e_2, t_1, t_2\}]$ .She knows she normally prefer choices e to the choices *t*. But a choice *t* (such as an extremely high possible payoff) may seem attractive from another perspective. Even though she is unsure about what will be her mood when making final choices, she thinks the s_u is more likely than s_v. Let us say $q(s_u) = 0.8$. Note that two options in $\{e_1, e_2\}$ and two options in $\{t_1, t_2\}$ have close riskiness levels. But r_u and r_v will determine a unique maximiser. For instance, if s_u is realised, e_1 is the maximum one with $u_{s_u}(e_1) = 1$ while $u_{s_u}(t_1) = 0$ and $u_{s_u}(t_2) = 0.01$; if s_v is realised, t_1 has the maximum utility with $u_{s_v}(t_1) = 0.9$ while $u_{s_u}(e_1) = 0.2$. So, the expected utility of menu $\{e_1, e_2\}$ becomes 0.8, that of $\{t_1, t_2\}$ is 0.72, however the menu $\{e_1, e_2, t_1, t_2\}$ has a utility of 1.52, and so $\{e_1, e_2, t_1, t_2\}$ will be preferred to other menus.

So, the main parameters of interest of PF are r_v and the state probability $q(s_v)$ [8].

5.2 Extensions

In our experiment, one menu of *f₃* and *f₄* contains two different temptations, gambling and risk-free lotteries. We are interested whether the presence of temptation diversity will influence the decision making. We assume there is a case that a DM is tempted by gambling and risk-free lotteries by the same time. Thus, the above utility functional of self-control and flexibility should be extended into more exogenous states. For the PF, the extension is straightforward. We assume there are three states, s_u the normative preference state where has a maximiser determined by r_u, s_{v1} the temptation state evoked by risk-free temptation where has a maximiser determined by r_{v1} (i.e, when the state happens, the risk-free lottery will be preferred) and s_{v_2} the temptation state evoked by gambling temptation where has a maximiser determined by r_{v2} (that is, when the state happens, the gambling lottery will be preferred). The probabilities of each state is measured by the $q(s)$ as equation (3). Therefore, the parameters of this case are r_{v1} , r_{v2} and correspoidng anticipated probabilies $q(s_{v_1})$ and $q(s_{v_2})$.

As for the SC, we can follow CK's extension to finite exogenous states, which is similar to the flexibility. We assume two states of world $S = \{s_{v_1}, s_{v_2}\}$ with the probability that the state occurs being given by $q(s)$. The state is realised after the DM chooses the menu. As CK's argument, DM's utility function does not change across states nor does her alter ego's. The only thing that changes is the probability of getting tempted. The menu utility functional can be written as

[8] Indeed, a more general method is to assume the belief on future preference contingency is continuous distribution, i.e., the anticipated risk preference is distributed normally with certain mean and standard deviation. The probability of choosing each available options will be determined by density distribution function of preference distribution. But it requires larger data sample.

$U_{SC}(A) = \sum_{s \in S} q(s)((1- p_s)\max_{x \in A} u(x) + p_s \max_{y \in B_v(A)} u(y))$ where p_s is the probability of getting tempted in the state of world.

The parameters here are r_{v1} , r_{v2}, the corresponding anticipated state probabilies $q(s_{v_1})$, $q(s_{v_2})$, and the probability of getting tempted in each state p_s.

5.3 Luce model

The choice *of* menu stage in all models are deterministic stories, identifying a particular optimal menu. In any experiment, however, there is behavioural noise. This fact implies that we have to model choices *of* menus in a stochastic fashion; otherwise, no model can explain the data. We use the multinomial logit model (or Luce model) to incorporate stochasticity. According to this model, the DM evaluates the problems with some noise. If the noise in the evaluation is additively separable and independently distributed according to the extreme value distribution, then the multinomial logit model emerges. This model implies that the *probability* of selecting one menu over another from a set of many menus is not affected by the presence or absence of other menus in the same context. The choice probability formula is given by the equation below.

$$P_i = \frac{e^{\lambda U_i}}{\sum_{j \in \Theta} e^{\lambda U_j}}$$

Here U_i is the expected utility of the menu *i*, *j* is any other menu in the menuset and *λ* is a precision parameter which measures the amount of experimental noise, and reflects the variance of the unobserved portion of utility.

5.4 Type identification

Decision rules are the private information of the DM. If instead, they were observable and verifiable by an outside party, one could simply contract upon them in a way given their decision rules. Indeed, decisions under different decision rules will differ in some cases. But as Table 1, indicates, sometimes different types' choice of menu could be identical. For example, within a limited sample, we may observe one DM keeping on choosing $\{e_1, e_2, t_1, t_2\}$ all the time. We cannot simply make a conclusion as to whether the decision is more likely to be coming from PF since a SC decision maker would randomly choose between two menus for which she is indifferent. The Luce model enables us to better identify the types based on maximum likelihood estimation. Essentially, the fundamental differences between each type are how subjects respond to temptation. Their attitude toward temptation will be incorporated into the menu utility. As disused, flexibility subjects respond to them by delaying commitment, thus the existence of temptation will increase the utility of the menu; self-control subjects respond to them by commitment themselves at choice of menu stage, thus the existence of temptation will create disutility, while the standard decision makers will not be influenced by the temptation. Whether the subjects are tempted, and are facing preference conflict, cannot be directly observed by the experimenter, but the choice probability distribution over menus is identifiable by our quantitative estimation. As long as subjects are tempted, the preference conflict will be measured by the parameters r_v and anticipated probabilities of Kreps and CK.

We assume that subjects are different. We therefore fit each of three preference functionals discussed above for each of the 82 subjects individually by maximum likelihood estimation. For every subject with the *n* observations of menu choices *i* from the menusets *S(f)* , recall that each menuset has three menus formed in particular way , the likelihood function can be written as

$LogL(\lambda, \Omega) = \sum_{n=1, i \in S(f)}^{N} \ln(P_{i,n})$ where Ω is the parameters in each menu utility functionals which will determine the choice probabilities $P_{i,n}$. With the same observation, the likelihood of different types will be different, as the $P_{i,n}$ is determined by utility functional *U*. We demonstrate identifiability with our Luce stochastic specification[9] through a simulation (see the Appendix).

6. Results

[9] Of course, this assumes that our subjects are noisy in their responses.

In this section, we present the experimental results from the experiment conducted at EXEC, the Centre for Experimental Economics at the University of York, in 2022. There were 81 subjects (mainly students from university of York) participating in the experiment. The mean earnings were £17.32 per subject (including a £2.50 show-up fee).

We started with the data from Part 1 of the experiment. There we effectively elicited the certainty equivalents of 28 lotteries. We assumed that the preferences of the subjects were Neumann-Morgenstern with an SMRA-CRRA[10] utility function with risk-aversion parameter r_u and we estimated the value of r_u for each subject. We assumed that noise (in the expressed certainty equivalents) was an additive normal distribution with mean zero and standard deviation 1/s. We also estimated *s* – the precision. The average estimated r_u is 0.04 with a standard deviation of 1.06 among subjects and the average estimated *s* is 0.04 with a standard deviation of 0.007 among subjects.

6.1 Overall menu frames effect

Our key fundamental research purpose is to understand whether different frames of menusets influence the choice of the menu and the choice from the menu. We start with some descriptive statistics.

Table 2 illustrates average individual choice *of* menu choice frequencies across menus under different frames. As shown in this table, f_2 significantly increases individual choice frequencies of menu $\{e_1,e_2,t_1,t_2\}$ compared to f_1 in both the gamble temptation and the risk-free temptation context (f_1 and f_2 with gamble temptation $p<0.05$; f_1 and f_2 with risk-free temptation $p<0.05$).f_3, Adding one risk-free temptation into the menu $\{t_1,t_2\}$ significantly increases the choice frequency of the pure temptation menu $\{t_1,t_2,t_3\}$ than $\{t_1,t_2\}$ in f_1 with both gamble ($p=0.00$) and risk-free ($p=0.05$)temptation cases. While the significance changes ($p<0.05$) are only observed between the choice frequencies of $\{e_2,t_2\}$ in f_2 with gamble temptation and $\{e_2, t_2,t_3\}$ *of* f_4. Overall, we can observe choice frequencies differences between *frames*.

[10] CRRA – Constant Relative Risk Averse; SMRA – Stochastically More Risk Averse (see Wilcox,2011)

	f_1			f_2		
	$\{e_1,e_2\}$	$\{t_1,t_2\}$	$\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$	$\{e_2,t_2\}$	$\{e_1,e_2,t_1,t_2\}$
gamble	0.53	0.09	0.38	0.24	0.10	0.66
risk	0.33	0.15	0.52	0.09	0.18	0.73
	f_3			f_4		
	$\{e_1,e_2\}$	$\{t_1,t_2,t_3\}$	$\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$	$\{e_2, t_2,t_3\}$	$\{e_1,e_2,t_1,t_2\}$
Multiple	0.45	0.22	0.32	0.31	0.19	0.50

All the choice frequencies of each subject are based on 10 observations.

Table 2: choice frequencies under different frames and temptations

For the choice *from* the menu, we are interested in the frequencies of choosing temptation choices under different frames. Table 3 shows the average individual choice frequencies of choosing temptation (that is, t_1,t_2,or t_3) as the final choice and choosing *ex ante* preferred choices (that is, {e_1 or e_2} Overall choice frequencies of choosing temptations are not significantly different among different frames even though the frequencies of choosing menus $\{e_1,e_2,t_1,t_2\}$ are higher under f_2 in the gamble temptation contexts. However, the frames have an effect on the frequencies of choosing risk-free temptation and multiple temptation options. Subjects have a significantly higher tendency to choose temptations in f_2 than in f_1. Note that all the temptation choices are designed with a lower expected payoff than any other lottery. To some extent, the gamble and risk-free temptations have different effects on choice frequencies of choosing temptation. Risk-free as temptation cases increase the frequencies of choosing temptation choices than gamble temptation cases under the same frame. There is strong evidence shows that the frequencies of choosing temptation with multiple temptations are higher than with gamble temptation.

	f_1		f_2	
	Temptation	*Ex ante* preferred choice	Temptation	*Ex ante* preferred choice
Gamble	0.13	0.87	0.15	0.85
risk-free	0.27	0.73	0.34	0.66
	f_3		f_4	
	Temptation	*Ex ante* preferred choice	Temptation	*Ex ante* preferred choice
Multiple	0.26	0.74	0.35	0.65

Table 3: Choice frequencies of choosing temptation under different frames and temptations

5.2 Type identification through Menu preference

Even though a frame effect has been observed, our two main postulates are that different types of subjects react to temptation differently; and different frames of menu have different effects on each type. We investigated whether any frames will increase the attractiveness of some *ex ante* undesirable choices. Thus, our analysis starts from *types* identification by fitting the choice *of* menu with three menu utility functionals. We take as given the estimated value of r_u, the risk-aversion of the *u* function. We estimated, by Maximum Likelihood, each of the 3 choice *of* menu functionals for the 81 subjects subject-by-subject, using the data on the choice *of* the menu, obtaining estimates of the parameters of the functional (particularly the implied CRRA level of risk-aversion r_u of the *v* function) and the maximised log-likelihood. In order to compare the relative goodness-of-fit of the mode, we need to correct the log-likelihoods for the different number of parameters (Kreps has 2, CK has 2 and Standard Theory has none). As the parameters of each function differ and the sample size for each subject is limited, we calculate the corrected likelihood value by the AICc[11] (Akaike Information Criterion). We define which type the subjects are more likely to be according to the model with lowest AIC.

Table 4 shows the fraction of each type under different frames with different temptations. Each *column* reports the *type* distribution under particular menu frames. Starting from the f_1 with gamble temptation subsets, 50% subjects are identified as PF type, 35% as SC type, and 15% as SD type. Comparing with the fraction of SC type in f_1, only 19% of subjects are identified as SD under f_2 with gamble temptations. The PF types have opposite changes. A larger fraction of subjects (68%) are identified as PF type under f_2 with gamble temptation. A similar pattern of the fraction of SD and PF types is observed between the f_1 and f_2 with risk-free temptation. We interpret this as *context-dependence menu preference*: subjects are more likely to respond to the temptation as PF when mixing the temptation with normatively preferred lotteries such as f_2, that is , $[\{e_1,t_1\},\{e_2,t_2\}\{e_1,e_2,t_1,t_2\}$.

Now the multiple temptations, adding risk-free temptation into one menu of f_1 and f_2 with gamble temptation, shows a different pattern. Recall that the f_3 is analogous to the menusets under f_1 with gamble temptations; and f_4 is analogous to that of f_2 with gamble temptations. If subjects will not be tempted by the two temptations, *risk-free and gambling lotteries* at the same time, adding one risk-free

[11] In small samples, AIC tends to over-fit. To address overfitting, AICc adds a size-dependent correction term that increases the penalty on the number of parameters (Burnham et al, 2002).

choices into menu will not make any difference. However, comparing the fraction under f_3 to the fraction under f_1 with gamble temptation, more subjects are identified as SD type (74% vs. 35%). Similarly, 97% of subjects are identified as SC under f_4 while only 19% under f_2 with gamble temptations. Subjects do feel tempted by two extreme temptations simultaneously. It implies subjects are more likely to react to temptation in a self-control way if the context evokes more conflicting preferences.

	f_1		f_2	
	gamble	risk-free	gamble	risk-free
PF	0.5	0.58	0.68	0.68
SC	0.35	0.10	0.19	0.02
ST	0.15	0.32	0.13	0.31
	f_3		f_4	
	Multiple temptation		Multiple temptation	
PF	0.26		0.03	
SC	0.74		0.97	
ST	0		0	

Table 4: Fraction of *types* under different frames with different temptations[12]

5.3 Frame effect on different types' decision-making

How does the menu frame influence different types' subjects' choice frequencies of choice of menus? In Table 5, we investigate each the *type* difference on choice *of* menu and choice *from* menu respectively. Comparing columns shows the observed choice frequencies differ across different types of subjects. It is consistent with the theory that PF types of subjects have a higher tendency to choose the flexibility menu $\{e_1, e_2, t_1, t_2\}$. Not surprisingly SC types tend to choose the menu $\{e_1, e_2\}$. Interestingly, the choice frequency distribution of SC types shows different patterns between f_1 and f_2. SC subjects have significantly ($p<0.05$) higher tendency to choose menu $\{e_1, e_2, t_1, t_2\}$ under f_2 than f_1 with gamble temptation cases. However, the significance is not observed in the risk-free cases. One of the main reasons is that only one subject is identified as SC.

[12] This table is constructed after identifying the best-fitting type of subject column by column – as described above; note that the entries in the columns add up to 1.

The PF *type* under f_3 with multiple temptations shows a higher tendency to choose *{t1,t2,t3}* in f_3 (51%) than with a similar menu *{t1,t2}* under f_1 (8%). Given the utility function of Kreps, the possible explanation is subjects feel tempted by risk-free and gamble temptations simultaneously and perceive a higher possibility to choose the risk-free temptations than the choice of normative preference, that is, e_1 *or* e_2 at the second stage. The ST type subjects show similar choice frequency pattern between f_1 (f_2) with gamble temptation and f_3 (f_4) with multiple temptations. When the commitment menu, *{e1,e2}* is available , they will tend to choose this menu to exclude the temptation.

		f_1			f_2		
		$\{e_1,e_2\}$	$\{t_1,t_2\}$	$\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$	$\{e_2,t_2\}$	$\{e_1,e_2,t_1,t_2\}$
Gamble	PF	0.36	0.08	0.56	0.26	0.09	0.65
	ST	0.59	0.21	0.20	0.24	0.24	0.53
	SC	0.74	0.05	0.21	0.18	0.03	0.79
Risk-free	PF	0.28	0.08	0.69	0.00	0.20	0.80
	ST	0.35	0.32	0.33	0.24	0.15	0.61
	SC	0.9	0	0.1	0.8	0	0.2[13]
		f_3			f_4		
		$\{e_1,e_2\}$	$\{t_1,t_2,t_3\}$	$\{e_1,e_2,t_1,t_2\}$	$\{e_1,t_1\}$	$\{e_2, t_2,t_3\}$	$\{e_1,e_2,t_1,t_2\}$
Multiple	PF	0.10	0.51	0.39	0.30	0.20	0.50
	ST	-	-	-	-	-	-
	SC	0.58	0.12	0.3	0.46	0.22	0.31

Table 5: Menu choice frequencies of each type under different frames and temptations

As for the choice *from* the menu, Table 6 shows each types' subjects' choice frequencies of choosing temptations and normal choices. Comparing PF and SC in f_1 with that of f_2, the frequencies of choosing temptation do not show a significant difference, while $f_2's$ ST subjects have a significantly increased frequency to choose temptation than $f_1's$ ST subjects for both menusets with gamble temptation and risk-free temptations.

Given the observation of PF type's higher tendency to choose menu *{t1,t2,t3}*, it is not surprising to observe the frequencies of choosing the temptation for PF type under f_3 with multiple temptation menu is significantly higher ($p<0.05$) than that under f_1. However, the significance cannot be found between

[13] Only 1 subject has been identified as a CK type.

f_2 and f_4. We can conclude that the PF subjects are more likely to give into the temptation in the multiple temptation frame.

		f_1		f_2	
		Temptation	*Ex ante* preferred	Temptation	*Ex ante* preferred
gamble	PF	0.14	0.86	0.13	0.87
	SC	0.05	0.95	0.2	0.8
	ST	0.27	0.73	0.21	0.78
Risk-free	PF	0.37	0.63	0.13	0.87
	SC	0.03	0.96	0.2	0.8
	ST	0.37	0.63	0.21	0.78
		f_3		f_4	
		Temptation	*Ex ante* preferred	Temptation	*Ex ante* preferred
Multiple	PF	0.55	0.45	0.3	0.70
	SC	0.16	0.84	0.26	0.74
	ST	-	-	-	-

Table 6: Choice from menu frequencies of each *type* under different frames and temptations

7. Conclusion

We report on an experimental investigation into two-stage decision-making, and investigate the influence of temptation and flexibility. The preference of menus has been studied extensively in self-control, commitment demand and flexibility theories and, to a much lesser extent, in experiments. However, we are unaware of any previous work that discusses the existence and effect of the desire for flexibility and self-control in the same context. This paper contributes to the small but growing literature applying insights and formal models from behavioural economics to the study of frame effects. In addition, our experimental design enables us to measure the importance of preference conflict, and to identify the possible behavioural principles of subjects under different frames of menusets. Designing the experiment to make the preference conflict identifiable, and behaviour patterns distinguishable, while following the basic structure of the models, was challenging, and we are aware of no other published experiment in which this behaviour has been demonstrated.

To identify the self-control, flexibility and standard theory types, the experiment was carefully designed. The identification strategies rely on the varied riskiness of lotteries to define the *ex ante* preferred choices and the *ex ante* unfavourable choices in menusets and the corresponding construction of menus. First, four frames of menusets are designed to identify different types of decision maker. We elicited each subject's *ex ante* risk preference in a menu-free context. Second, we place the *ex ante* preferred lotteries in different menus with *ex ante* unfavourable lotteries (which we term as gambling temptations and risk-free temptations). With the design of lotteries, menu frames and the application of Luce model, the directly unobservable preference conflict can be captured.

Our menu designs are not just for econometric convenience, but also offer real world business insight. Our principal conclusion can be summarised as follows. First, we show that f_2 $[\{e_1,t_1\},\{e_2,t_2\},\{e_1,e_2,t_1,t_2\}]$, which mixed the normatively preferred choices with *ex ante* undesirable choices can significantly increase the choice of menu $\{e_1,e_2,t_1,t_2\}$. However, it does not influence the final choice from the menu: most subjects still choose from the menu according to their *ex ante* normative preferences.

Secondly, the frame does have a significantly adverse impact on the self-control type subjects. As the theory predicts, subjects with self-control problems have stronger commitment demand. Removing the commitment menu will not only lead to higher possibilities of choosing the menu $\{e_1,e_2,t_1,t_2\}$, but also, self-control subjects have a higher tendency to give into the temptation. The self-control subjects' average payoffs of the chosen lotteries are all lower in f_2 , for example subjects who are identified as self-control subjects have an average payoff of £11.18 under f_1 while £9.53 under f_2 with gamble temptation context; and earned £13.69 under f_1 while £11.4 under f_2 with risk-free temptation averagely. Offering menu choices without commitment menu lead self-control subjects to perform worse.

Third, f_3 and f_4 (adding a risk-free temptation into a menu with gamble temptation, which we term as higher flexibility) shows a different impact on self-control subjects than on flexibility subjects. Multiple temptations create more preference conflicts, and will lead more subjects to exclude the presence of temptations. However, increasing the flexibility has stronger effect on the PF types. This kind of behaviour is common in real life: when choosing an insurance scheme, people are more likely to be convinced if there is some minimum guaranteed amount; when choosing a risky investment portfolio, people are more likely to be convinced if they are told the minimum return. It gives some insight to companies wanting to launch a new product, or to policy makers who want consumers to accept more risk, by offering safe options in a risky bundle.

Our results clearly show the three different models' explanatory power concerning preference inconsistency. A considerable of research in this line has applied the theories to time inconsistency. Following Gul and Pesendorfer (2004), Krusell and Smith, Jr., (2003), Laibson *et al*, (1998), and using our experimental design, we can extend the choice context into other dynamic decisions, such as long-term consumption and savings. For example, Gul and Pesendorfer's model using the self-control story formalises time-inconsistent behaviour driven by the conflict between the temptation of smaller-earlier rewards and larger-later rewards. Our temptation design with risk-free options can be implemented as smaller-earlier (which can be played out immediately at each stage), while the *ex ante* risk preference and very-risky option can be implemented as larger-later reward. Offering subjects different rewards with three levels of riskiness (that is, risk-free, *ex ante risk* preferred, and very-risky) enables us to quantitatively and parametrically measure how subjects resolve the intertemporal trade-off between immediate rewards, larger-later rewards and gambling across stages in life-cycle consumption models. We expect future research using our experimental design to explore further self-control and temptation models in a long-term dynamic context. We can also investigate optimal contract design (menu design) from observations of self-control behaviour in this context.

References

Ahn, D. and Sarver, T., 2013. "Preference for Flexibility and Random Choice". *Econometrica*, 81(1), pp.341-361.

Amador, M., Werning, I., & Angeletos, G. (2006). Commitment vs. Flexibility. *Econometrica*, *74*(2), 365-396. doi: 10.1111/j.1468-0262.2006.00666.x

Bauer, P., & König, F. (2014). "The risks of methodology aversion in drug regulation". *Nature Reviews Drug Discovery*, 13(5), 317-318. doi: 10.1038/nrd4319

Bond, P., & Sigurdsson, G. (2013). Commitment Contracts. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2329834

Coudert, V., & Gex, M. (2008). "Does risk aversion drive financial crises? Testing the predictive power of empirical indicators". *Journal of Empirical Finance*, 15(2), 167-184. doi: 10.1016/j.jempfin.2007.06.001

Chatterjee, K. and Krishna, R., 2009. "A 'Dual Self' Representation for Stochastic Temptation". *American Economic Journal: Microeconomics*, 1(2), pp.148-167.

Dekel, E. and Lipman, B., 2012."Costly Self-Control and Random Self-Indulgence". *Econometrica*, 80(3), pp.1271-1302.

Dekel, E., Lipman B. and Rustichini, A., 2009. "Temptation-Driven Preferences". *Review of Economic Studies*, 76(3), pp.937-971.

Dekel, E., Lipman, B., & Rustichini, A. (2001). Representing Preferences with a Unique Subjective State Space. *Econometrica*, *69*(4), 891-934. doi: 10.1111/1468-0262.00224

DellaVigna, S., & Malmendier, U. (2004). "Contract Design and Self-Control: Theory and Evidence". The *Quarterly Journal of Economics*, 119(2), 353-402. doi: 10.1162/0033553041382111

Esteban, S., & Miyagawa, E. (2022). Optimal menu of menus with self-control preferences. Retrieved 28 June 2022, from https://academiccommons.columbia.edu/doi/10.7916/D8V69WSB

Gul, F. and Pesendorfer, W., 2001. "Temptation and Self-Control". *Econometrica*, 69(6), pp.1403-1435.

Houser, D., Schunk, D., Winter, J. and Xiao, E., 2010. "Temptation and Commitment in the Laboratory". *SSRN Electronic Journal*

Hyeongmin (Christian) Kim, "Situational Materialism: How Entering Lotteries May Undermine Self-Control", *Journal of Consumer Research*, Volume 40, Issue 4, 1 December 2013, Pages 759–772, https://doi.org/10.1086/673191

Lin, L. (2021). "Does the procedure matter?". *Journal of Behavioral and Experimental Economics*, 90, 101618. doi: 10.1016/j.socec.2020.101618

Laibson, D., Repetto, A., Tobacman, J., Hall, R., Gale, W., & Akerlof, G. (1998). Self-Control and Saving for Retirement. *Brookings Papers On Economic Activity*, *1998*(1), 91. doi: 10.2307/2534671

Koida, N., 2017. "Anticipated stochastic choice". *Economic Theory*, 65(3), pp.545-574.

Kreps, D., 1979. "A Representation Theorem for 'Preference for Flexibility' " *Econometrica*, 47(3), p.565.

Krusell, P., & Smith, Jr., A. (2003). Consumption-Savings Decisions with Quasi-Geometric Discounting. *Econometrica*, *71*(1), 365-375. doi: 10.1111/1468-0262.00400

Masatlioglu, Y., Nakajima, D., & Ozdenoren, E. (2020). "Willpower and compromise effect". *Theoretical Economics*, 15(1), 279-317. doi: 10.3982/te3086

Noor, J., & Takeoka, N. (2015). "Menu-dependent self-control". *Journal of Mathematical Economics*, 61, 1-20. doi: 10.1016/j.jmateco.2015.07.004

Ortoleva, P., 2013. "The price of flexibility: Towards a theory of Thinking Aversion". *Journal of Economic Theory*, 148(3), pp.903-934.

Stovall, J., 2010. "Multiple Temptations". *Econometrica*, 78(1), pp.349-376.

Salant, Y., & Siegel, R. (2018). Contracts with Framing. *American Economic Journal: Microeconomics*, *10*(3), 315-346. doi: 10.1257/mic.20160230

Toussaert, S., 2018. "Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment", *Econometrica*, 86(3), pp.859-889.

Wilcox, N. (2011), "Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk*", Journal of Econometrics,* 162, 89–104

Appendix A

In this simulation we have assumed that r_U=1.73 and have used the corresponding menusets generated as in our experiment as described above. We first generated 100 sets of observations on the choice *of* a menu under frame f_1 with gamble temptations for the different models assuming Luce noise. We then estimated the parameters of the different models. This was to see if the maximum likelihood estimation can identify the true model that was used to generate the decisions, and if the true parameters can be estimated. We use the AICc criterion to determine the best-fitting model.

True model is SD with r_u=1.73			
Models	r_v	p	Corrected likelihood
Kreps	2.6	0	11.99
CK	2.4	0.16	12.03
SD	-	-	11.98

Table A1: Average estimated parameters and corrected likelihood value (AICc)

True model is Kreps with r_u=1.73, r_v=-2, p=0.8			
Models	r_v	p	Corrected likelihood
Kreps	-2.23	0.85	13.9
CK	-3	0.98	15.27
SD	-	-	15.44

Table A2 Average estimated parameters and corrected likelihood value (AICc)

True model is CK with r_u=1.73, r_v=-2, p=0.8			
Models	r_v	p	Corrected likelihood
Kreps	-5	0.95	14.15
CK	-2.4	0.79	1.95
SD	-	-	18.22

Table A2 Average estimated parameters and corrected likelihood value (AICc)

The above tables show that using the different menu utility functional with Luce model are tractable and identifiable.