Preference for Flexibility and Random Choice: an Experimental Analysis*

Mark Dean and John McNeill

August 18, 2014

Abstract

People may be uncertain about future preferences, leading to both a preference for flexibility in choice between menus and stochastic choice from menus. This paper describes an experimental test of preference uncertainty in a real-effort task. We observe subjects' preferences over menus of work contracts, along with their choices of effort levels from those contracts. Our results suggest that preference uncertainty is important: 48% of our subjects exhibited strict preference for flexibility. A model of preference uncertainty (Ahn and Sarver (2013)) well describes the relationship between choice of and from menus: subjects willing to pay to include an option in a contact were more likely to use that option, and those that used an option were prepared to pay for it. We show that the introduction of an explicit stochastic element to the contract increased preference for flexibility, suggesting a causal role for uncertainty in menu preferences.[1]

1 Introduction

Preferences may be both unstable and unpredictable. Some days you may be full of energy and willing to put in a lot of work in return for more pay. On others, you

*The authors gratefully acknowledge support from NSF Grant NSF 1156090, and thank Marina Agranov, Pedro dal Bó, Pietro Ortoleva and participants at the Brown Theory Lunch for helpful suggestions.

[1]This experiment received IRB approval from Brown University (Protocol #1301000749) on March 05, 2013.

may feel lazy and sluggish and be much less willing to work hard. Moreover, it may be difficult to predict in advance in which of these states you will find yourself on a particular day.

One implication of this observation is that people may have a preference for flexibility. If a decision maker (DM) is uncertain about his or her future cost of effort he or she may want to have a number of different output/payment combinations available to choose from when the uncertainty resolves itself. This can lead to a strict preference for larger menus (i.e. sets of options from which a choice will later be made). A second implication is that choices may be stochastic: if a DM's preferences vary then so will his or her subsequent choices.

This paper experimentally explores the extent to which preference uncertainty is an important component of choice behavior in a real-effort task. It has three specific aims. First, to measure the extent to which workers exhibit preference for flexibility. Second, to test whether preference over menus is linked to choice from menus in the manner predicted by Ahn and Sarver (2013),[2] who provide a unified account of these phenomena by giving conditions under which the same underlying uncertainty over preferences can explain both behaviors. Finally, to examine the effect of introducing an explicit stochastic element to worker contracts in order to establish a causal link between uncertainty and preference for flexibility.

Subjects in our experiment were offered the chance to perform simple arithmetic tasks for money. The experiment was run using Amazon's Mechanical Turk platform, an Internet marketplace through which people are hired to perform small tasks that are not easily automated, such as podcast transcribing or image tagging. A significant advantage of using the Mechanical Turk platform is the close relationship between the tasks involved in our experiment and the tasks that these workers usually perform, meaning that we classify our study as a 'quasi-field' experiment.

Subjects' payments were governed by an employment contract which described how much the subject would earn conditional on achieving a target number of tasks. A 'Low' contract had a low target and low payment, a 'High' contract a high target and high payment, and a 'Flex' contract both low and high targets, with the corresponding payments. Subjects first worked under exogenously specified contracts, before being offered the chance to choose between different contracts for subsequent work sessions.

[2]This model combines the preference for flexibility model of Dekel et al. (2007) with the stochastic choice model of Gul and Pesendorfer (2006).

Our first result is that subjects had a significant degree of preference for flexibility. 35% of well-behaved subjects[3] exhibited a strict preference for flexibility (i.e. were prepared to pay for the Flex contract over both the Low and High contracts) in our baseline specification, and 48% exhibited such preferences for some parameter values. Moreover, 94% of subjects demonstrated preferences in line with set monotonicity (i.e. never strictly prefer a smaller set), a necessary condition for preference uncertainty to explain menu preferences (Dekel, Lipman and Rustichini (2001); Ahn and Sarver (2013)).

Our second result is that the subjects' choices from menus were largely consistent with the same underlying uncertainty revealed by choice from menus, in line with the model of Ahn and Sarver (2013). We estimated stochastic choice functions for our population as a whole, and for subsets of subjects categorized by their preference between menus. Ahn and Sarver (2013) provides two conditions that are necessary to view preference for flexibility and stochastic choice as being driven by the same underlying preference uncertainty. The first is that a DM who is prepared to pay for an alternative to be included in his or her menu must use that alternative some of the time. The second is that a DM who uses an alternative must be prepared to pay to have it in his or her menu (except in the case of indifference). We find evidence for both of those effects: subjects who paid for the Flex contract over the Low contract were significantly more likely to use the high target, while those who paid for the Flex contract over the High contract were significantly more likely to use the low target. As a result, subjects with preference for flexibility made substantial use of both contracts. Moreover, 96% of non-indifferent subjects who used the low target strictly preferred the Flex contract to the High contract, while 83% of such subjects who used the high target strictly preferred the Flex contract to the Low contract.

Our final result is that an exogenous increase in uncertainty leads to an increase in preference for flexibility. We demonstrate this by running a further experiment which introduced a stochastic component to the number of tasks that needed to be completed for the High target. This uncertainty was resolved after the contract was chosen but before it was implemented. Thus, subjects did not know how many tasks they would have to complete for the High target at the time that they choose

[3]We describe what we mean by well-behaved in section 4.1 - essentially it means subjects who avoided making dominated choices. Due to the noisy nature of Mechanical Turk we have a significant number of subjects - 48% - who failed this test. However, including these subjects actually increases the proportion who exhibited preference for flexibility in the baseline specification to 43%.

the contract, but did know when the contract was in effect. We find an increased preference for flexibility under such a 'random' contract. This result suggests that uncertainty plays a causal role in determining menu preferences.

It is of interest to try to understand what it is that our subjects were uncertain about in the baseline (deterministic) experiment. Our favored interpretation is that it was their relative valuation of effort and monetary reward at the time they must perform the task. In Section 6 we argue against an alternative explanation - that subjects were uncertain about their ability to complete the task within the allotted time. Only a small number of subjects were in danger of running out of time, and those subjects were somewhat less likely to exhibit preference for flexibility.

To our knowledge our paper is the first to experimentally document and study preference for flexibility in menu choice.[4] We are also the first to link such preferences to subsequent choice from menus. This link helps to identify preference uncertainty as the driving force for preference for larger menus, a mechanism that has been posited since at least Kreps (1979), but never tested. Our work also has relevance for the large theoretical and practical literature relating to temptation and self-control. Sophisticated agents who suffer from self-control problems may exhibit preference for commitment (i.e. a preference for smaller choice sets), yet there is (perhaps surprisingly) limited demand for commitment devices (though see recent work by Kaur, Kremer and Mullainathan (2010)). A high degree of preference uncertainty could be one possible explanation for this. As pointed out by Amador, Werning and Angeletos (2006), preference uncertainty could provide an offsetting preference for flexibility, making commitment costly. Our results suggest that it may be important to take into account preference uncertainty when designing employment contracts.

The remainder of the paper is organized as follows. Section 2 describes the theoretical background. Section 3 outlines the experimental design, 4 our data and identification strategies and 5 our results. Section 6 discusses the source of subject uncertainty. Section 7 describes the related literature, and 8 concludes.

[4]By 'flexibility' we mean that the union of two menus is preferred to either of the underlying menus. Other work (e.g. Sonsino and Mandelbaum (2001)) has documented the fact that people may pay for larger menus in line with standard theory.

2 Model

Our analysis is based around a model of behavior in which a decision maker faces uncertainty about his or her future preferences - in this case the intrinsic cost of effort. This can lead to both a preference for flexibility (in the manner of Dekel et al. (2007), henceforth DLRS) and stochastic choice (in the manner of Gul and Pesendorfer (2006), henceforth GP). Ahn and Sarver (2013) (henceforth AS) provides conditions under which the same uncertainty can be seen as driving both preference for flexibility and stochastic choice.

The building blocks of the 'preference uncertainty' model described by AS are a set of alternatives Z, lotteries over these alternatives $\Delta(Z)$, and menus of such lotteries A, with $\mathcal{A}$ denoting the set of such menus. The DM has preferences over $\Delta(Z)$ which depend on the (unobserved by the researcher) realization of a subjective state $s \in S$, and are modeled via a state dependent utility function $U : S \times \Delta(Z) \to \mathbb{R}$. The DM has beliefs about the likelihood of states denoted by the probability distribution $\mu \in \Delta(S)$.

This model can be used to understand the DM's preference over menus (which we represent with the complete preference relation $\succeq$ on $\mathcal{A}$). The assumption is that uncertainty about S resolves itself after the DM chooses a menu, but before he or she chooses from that menu. Thus menu preferences are represented by a utility function defined for every A by:

$$U(A) = \sum_{S} \mu(s) \cdot \max_{x \in A} U_s(x) \tag{1}$$

In words, 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.

The same model can also be used to describe the distribution of choices from a given menu. Using $\lambda : \mathcal{A} \to \Delta(\Delta(Z))$ to denote the probability of choosing each alternative from each menu, the model predicts that[5]

$$\lambda^A(x) = \sum_{s \in S} \mu(s) \mathbf{1}[x \in \operatorname*{argmax}_{x \in A}(U_s(x))] \tag{2}$$

The probability that an option x is chosen from a set A is equal to the probability

[5]Subject to the tie-breaking rule τ, which GP also address as another potential source of randomness.

of the states in which x is the highest utility object in A, as captured by the indicator function in Equation 2.

AS provides conditions under which a DM's preferences over menus are consistent with his or her (stochastic) choice from menus: in other words, the same beliefs and state-dependent utility function can be used to represent a DM's preferences over menus (in the sense of Equation 1) and stochastic choice function (in the sense of Equation 2).

Of particular interest to our study are four behavioral implications of the preference uncertainty model. The first is related to choice over menus, the second to choice from menus, and the third and fourth to the relationship between these two behaviors:[6]

1. **Preference for Flexibility:** Consider two menus A and B such that $A \succeq B$. If there were no uncertainty over future preferences (for example if S were a singleton) then Equation 1 would imply $A \cup B \sim A$. Yet if S is not a singleton, it is possible that $A \cup B \succ A$. This is what is generally referred to as strict 'Preference for Flexibility', and it results directly from preference uncertainty. Moreover, the preference uncertainty model rules out a strict preference for smaller choice sets, so $A \subseteq B \Longrightarrow B \succeq A$. It is this property which distinguishes the preference uncertainty models from models of temptation and self-control (for example Gul and Pesendorfer (2001)) in which smaller choice sets may be preferred to avoid exposure to tempting options.

2. **Choice Monotonicity:** When applied to choice from menus, the preference

[6]The AS representation requires all of the axioms used in the construction of the DLRS and GP representations, some of which we do not seek to test in our experiment. DLRS requires that preferences over menus be a weak order (DLRS1) and nontrivial (DLRS3) - we did not test these directly, but do discuss the fraction of subjects whose choices were intransitive and those who displayed strict preferences over menus. Implicitly we test the axioms of AS conditional on menu preferences being a weak order. We did not test the requirement of Lipschitz continuity on preferences over menus (DLRS2), or continuity of random choice rules in mixtures over decision problems (GP 'mixture continuity'). DLRS also requires 'weak independence' as an independence property on menu preferences, and although we do observe preferences over submenus, we did not mix them with other menus in a way that allows us to address failures of weak independence; similarly GP requires a form of independence ('linearity') that we did not address with this experimental framework, as we did not observe choices over lotteries - the uncertainty in our Random Contracts treatment was resolved before subjects chose from the menu. GP's 'extreme' property requires that the random choice rule selects a maximizer of some utility function with probability 1. In our set up, this is equivalent to never choosing a dominated number of tasks to complete. See Section 4.2.1 for a discussion of this issue.

uncertainty model is essentially a random utility model, and shares the implication of choice monotonicity: adding options to a choice set cannot increase the probability that an existing option will be chosen - i.e. $x \in A \subset A' \implies \lambda^{A'}(x) \leq \lambda^{A}(x)$. This is in contrast to attention-based models of stochastic choice (for example Caplin and Dean (2014)), in which new alternatives can alter attention in such a way to increase the likelihood of choosing a previously available option.

3. **Consequentialism:**[7] Consider a DM who strictly prefers adding an alternative x to menu A. According to Equation 1, this can only be the case if x has higher utility than all the elements of A in some state that occurs with non-zero probability. Thus if the same preference uncertainty is to explain stochastic choice, it must be that alternative x is chosen from the menu $A \cup \{x\}$ in some state. This is this empirical content of the consequentialism axiom of AS:

$$A \cup \{x\} \succ A \implies \lambda^{A \cup \{x\}}(x) > 0$$

4. **Responsive Menu Preferences:**[8] The natural counterpart to the consequentialism condition is that if x is sometimes chosen from the menu $A \cup \{x\}$, then this menu should be strictly preferred to A. This is, however, too strong, as it could be the case that the DM is indifferent between x and another element of A in the state in which it was chosen, and its selection was as a result of a tie-breaking rule. Thus, the fourth condition is that $\lambda^{A \cup \{x\}}(x) > 0$ implies $A \cup \{x\} \succ A$ except in cases of indifference. AS use continuity conditions to behaviorally rule out indifference. We discuss how we deal with this issue in section 5.2.3.

3 Experimental Design

Subjects were offered the opportunity to complete effort tasks for payment. Thus, final choices were made over effort/money pairs. Subjects chose from menus of such pairs - i.e. contracts that specified payment contingent on the number of tasks completed. We observed both subjects' choice of effort/money pairs from different contracts and

[7] This is Axiom 1 of AS.
[8] This is Axiom 2 of AS.

their preferences between contracts, allowing us to test the predictions discussed in the previous section. Implicitly, any preference uncertainty would relate to the relative value of effort to money at the point at which the task was to be completed.

The experiment was conducted using Amazon's Mechanical Turk (MT) platform. Mechanical Turk is digital marketplace for work; "requesters" post Human Intelligence Tasks, or "HITs", which are usually simple, repetitive tasks that are not easily automated, and also typically pay very small sums for each task. Workers on MT view descriptions of the HITs, decide whether or not to accept them, and complete accepted HITs over the Internet. The results of the HIT are submitted to the requester, who approves or rejects the HIT. In our case, subjects who accepted the HIT followed a link to an external webpage, where they completed the experiment. Upon completion they were given a randomly generated code, which was used to pay them the appropriate amount given their choices in the experiment.[9]

Mechanical Turk is a relatively novel environment in which to conduct economic experiments.[10] One key advantage for the current study is that it represents a 'quasi-field' setting: subjects were people who had signed up through the MT platform to supply labor in exchange for money in precisely the way that was on offer in our experiment. Thus for this study population the experimental setting is arguably more natural (and possibly more externally valid) than for the typical undergraduate pool.

Further advantages to using MT are the ready availability of subjects and the low cost of collecting data. The prevailing wage rate on MT is extremely low. The payments involved in our experiment were therefore low by the standard of traditional on-campus laboratory experiments, but were commensurate with the prevailing wages on MT. One cost of using MT rather than a traditional laboratory setting is a reduction in experimenter control: because MT workers complete the experiment remotely,

[9]The pool of workers on Mechanical Turk is global and diverse. It is possible to filter who is able to accept and complete a HIT according to different criteria, notably according to geographic location and the lifetime HIT approval rating for the worker (on all HITs completed, not just those related to this task). As we had no a priori reason to restrict our workers to a specific geographic subset, we imposed only the requirement that the worker's HIT approval rating must be above 90%, a requirement commonly found on MT. We did not target any particular demographics of MT workers for our recruitment. The HIT was described as a "decision-making experiment", and used the keywords "decision", "experiment", "study", "bonus", and "payment".

Some workers accepted the HIT but did not complete the experiment - in the vast majority of those cases (81%), the worker exited the experiment prior to completing the instructions.

[10]For more discussion about Amazon's Mechanical Turk, its mechanics, and its use in experimental social sciences, see Paolacci, Chandler and Stern (2010), Mason and Suri (2012), and Goodman, Cryder and Cheema (2013).

the experimental environment is not as tightly controlled as it would be in the laboratory. Accordingly it is important to pay particular attention to data quality, as subject are potentially less focused on the task at hand than they would be in the laboratory. We address this issue in section 4.1.

3.1 Main treatment

The four primary components of the experiment were the tasks that subjects could choose to complete to earn additional payment, the 'task sections' in which these tasks were performed, the contracts that specified how many tasks a subject had to complete in order to earn different payments in a task section, and multiple price list questions designed to elicit subject preferences over contracts.

3.1.1 Tasks

The building block of the experiment was a real-effort task: a simple activity that the subject could choose to perform numerous times in order to earn additional payment. Each task was an arithmetic problem that required the subject to add together two three-digit numbers. For example, subjects might have seen "369+458", and been required to enter "827" into a text box as the solution in order to complete the task. After submitting their proposed solution to the problem the subject was told whether the answer they submitted was correct or incorrect - correct answers counted towards his or her current total, while incorrect answers did not. Figure 1 shows an example of a typical task screen.

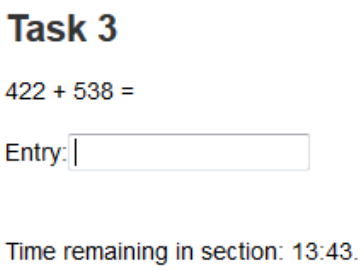


Figure 1: Example task

3.1.2 Task Sections

Subjects faced several 'task sections': opportunities to perform tasks in exchange for money. Payment in each task section was governed by a contract which specified how many tasks needed to be completed in order to earn specified additional payments (see Section 3.1.3). Subjects could complete as many or as few tasks as they wished, but had to do so within a time limit.[11] Once they began working on a given task section, subjects had 15 minutes to complete as many tasks as they wished, although they could "retire" from the section and continue on with the experiment at any time. Thus, a task section ended whenever the timer ran out or the subject decided to stop. Subjects could not return to a previous task section once the time had run out, or after they had chosen to retire from that section. For the contracts used in this experiment there was no benefit to subjects for completing more than 50 tasks, although there was no strict upper limit on the number of tasks they could potentially complete. This means that subjects who wished to earn the highest possible reward needed to complete on average at least one task every 18 seconds. The intention of this time limit was to keep the subjects focused on the task at hand, rather than to make it impossible or even difficult to complete the desired level of tasks. For most subjects, 18 seconds was ample time to complete a three-digit addition task with a moderate application of effort, but the 15 minute time limit on the section helped to discourage subjects from pausing in the middle of a section to do other things. In practice, subjects did not seem to encounter problems with the time limit, as we discuss in Section 6.

3.1.3 Contracts:

In each task section payment was governed by a contract which specified how many tasks must be completed in that section in order to receive an additional amount of payment, as shown in the example contracts shown in Figure 2.

Under Contract 24 (Figure 2), for example, completing fewer than 20 tasks would earn no additional compensation, completing between 20 and 49 tasks would earn \$0.20, and completing 50 or more tasks would earn a total of \$0.40 for that section. Note that this was not a piece-rate arrangement, so the amount specified was the

[11] While completing tasks, subjects were shown the number of tasks they had completed correctly in that section, as well as the payment that that number of tasks had earned.

Contract 11

Tasks completed	Payment
0-4	0.00
5-9	0.00
10-14	0.00
15-19	0.00
20-49	0.20
50+	0.20

Contract 25

Tasks completed	Payment
0-4	0.00
5-9	0.00
10-14	0.00
15-19	0.00
20-49	0.00
50+	0.40

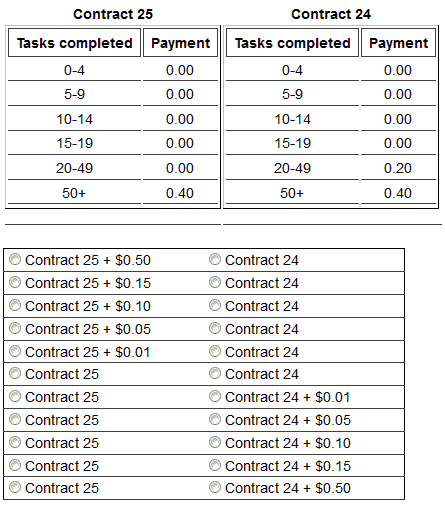
Contract 24

Tasks completed	Payment
0-4	0.00
5-9	0.00
10-14	0.00
15-19	0.00
20-49	0.20
50+	0.40

Figure 2: Low, High, and Flex contracts

total payment earned for completing a given number of tasks. It would be possible to achieve something similar using a piece-rate wage, in the spirit of Kaur, Kremer and Mullainathan (2010). However, the approach taken here has the advantage of producing a relatively small set of sensible choices the subjects can make: given that effort is costly, subjects should either choose to do no tasks at all or to do enough tasks to reach a given payment level and then stop. This has the benefit of resulting (in principle) in data in which subjects' efforts are clustered at one of a small number of different completion levels, so that their actions can be condensed into one of a few types.

Each contract used in the experiment had either one or two levels at which the payment increased. In each task section, subjects could complete as many or as few tasks as they chose, which always included the option to complete no tasks for no additional payment. The other level(s) in a given contract were either a low number of tasks for a lower payment or a high number of tasks for a higher payment. The contracts used in the main experimental treatment were of three types. The "High" contract included only the option to do 50 tasks for \$0.40 (Contract 25 in of Figure 2). The "Low" contracts offered only the option to receive \$0.20 for completing the low number of tasks. These contracts are indexed by the number of tasks required to earn the \$0.20. Our main analysis focuses on the case in which 20 tasks were required to receive the low payment (Contract 11 of Figure 2), which we will refer to as "Low-20" or "L20"; the contracts that subjects experienced in the exogenous contracts sections required 20 tasks to earn the low payment. However we also study the effect of changing the number of tasks required to receive the low payment; a "Low" contract requiring X tasks will be referred to as "LX". The "Flex-X" (or FX) contract combined the options available in the "High" contract and the "LX" contract, containing both the option to do X tasks for \$0.20 and the option to do the 50 tasks



Contract 25	
Tasks completed	Payment
0-4	0.00
5-9	0.00
10-14	0.00
15-19	0.00
20-49	0.00
50+	0.40

Contract 24	
Tasks completed	Payment
0-4	0.00
5-9	0.00
10-14	0.00
15-19	0.00
20-49	0.20
50+	0.40

Contract 25 + $0.50	Contract 24
Contract 25 + $0.15	Contract 24
Contract 25 + $0.10	Contract 24
Contract 25 + $0.05	Contract 24
Contract 25 + $0.01	Contract 24
Contract 25	Contract 24
Contract 25	Contract 24 + $0.01
Contract 25	Contract 24 + $0.05
Contract 25	Contract 24 + $0.10
Contract 25	Contract 24 + $0.15
Contract 25	Contract 24 + $0.50

Figure 3: Example question: H vs $F20$

for $0.40 (F20 is shown as Contract 24 in Figure 2).

3.1.4 Contract Preference Questions

The fourth component of the experiment was a set of questions used to elicit subjects' preferences over different contracts. Subjects responded to a series of multiple price list (MPL) questions which asked them to choose between two different contracts for potential use in a future task section, as well as side payments, as in Figure 3. Following these questions, one line from one question was selected at random, and the subject's choice on that line was implemented as the contract they used to complete a subsequent task section.

In each MPL question the subject made a series of eleven pairwise choices between the two options on each line. On each subsequent line the option on the right became more attractive relative to the option on the left, either because the side-payment associated with the left-hand option decreased or because the side payment associated with the right-hand option increased. This means that if the subject chose the right-hand option on a given line, they should also have chosen the right-hand option on each subsequent line, which provides a basic check that the subject was paying attention to his or her choices. Any subject who switched from right-to-left either had preferences that are not monotonically increasing in money, or (more probably) was not paying attention to the question. In all cases, the side payments were as

shown in Figure 3.

Typical responses to the MPL questions begin on the left and at some point switch over to the right. The point at which the subject switched translates to the strength of preference for the contract on the left compared to the contract on the right - a subject who was willing to forgo an additional payment (for example, choosing "Contract 24" instead of "Contract 25 + \$0.01") must have strictly preferred the contract selected. A subject who was indifferent between the contract on the left-hand side and the contract on the right-hand side would always select the option with additional payment, and could have chosen either option in the center line which has no additional payments, meaning they would have first chosen the right-hand side option on line six or seven. Switching over to the right before the sixth line, or after the seventh, demonstrates a strict preference for one contract over the other.

3.1.5 Experimental Structure:

The experiment began with a set of instructions that introduced the tasks and contracts, including a comprehension quiz that required subjects to demonstrate that they understood how to read the contracts and determine how much payment would accompany a given number of tasks completed.[12] Following the instructions and comprehension quiz, subjects were required to complete four practice tasks, to ensure a baseline level of familiarity with the experimental interface before they began the main part of the experiment. Following this, they completed three task sections. In each of these sections the subject was exogenously assigned a different contract. The three contracts used are those shown in Figure 2: L20, H and F20. The order of contracts was randomized between subjects.

Following the three exogenously specified contract sections, subjects responded to a set of nine questions to elicit their preferences over contracts to be used in a fourth and final task section. Subjects were presented with MPLs covering bilateral comparisons between H, L20 and F20. These questions refer to the experienced contracts, and are the focus of our main analysis. In addition, subjects answered MPL questions between H and FX and between LX and FX for X=15, 10 and 5. The order of the nine questions was randomized across subjects.

After the subject completed the block of questions, one of the lines of one of the questions was randomly selected to be realized in a fourth and final task section.

[12] A copy of the instructions is shown in Appendix B.

Note that the time lag between the choice of contract and the implementation of that contract was relatively short. We discuss this issue further in Section 6.

Following the final task section, the subject was paid according to his or her performance during the experiment through the MT interface for payments. Subjects were paid for their performance in all four task sections, as well as for the selected line in the MPL questions. Subjects also received a participation fee of $0.25.

3.2 "Random Contracts" Treatment

Any preference for flexibility found in the main treatment is naturally-occurring. In order to establish a causal link between uncertainty and preference for flexibility we ran a further experimental treatment which introduced exogenous uncertainty about some contract features, with the uncertainty resolved after the choice between menus but prior to the task section in which the contract was implemented.

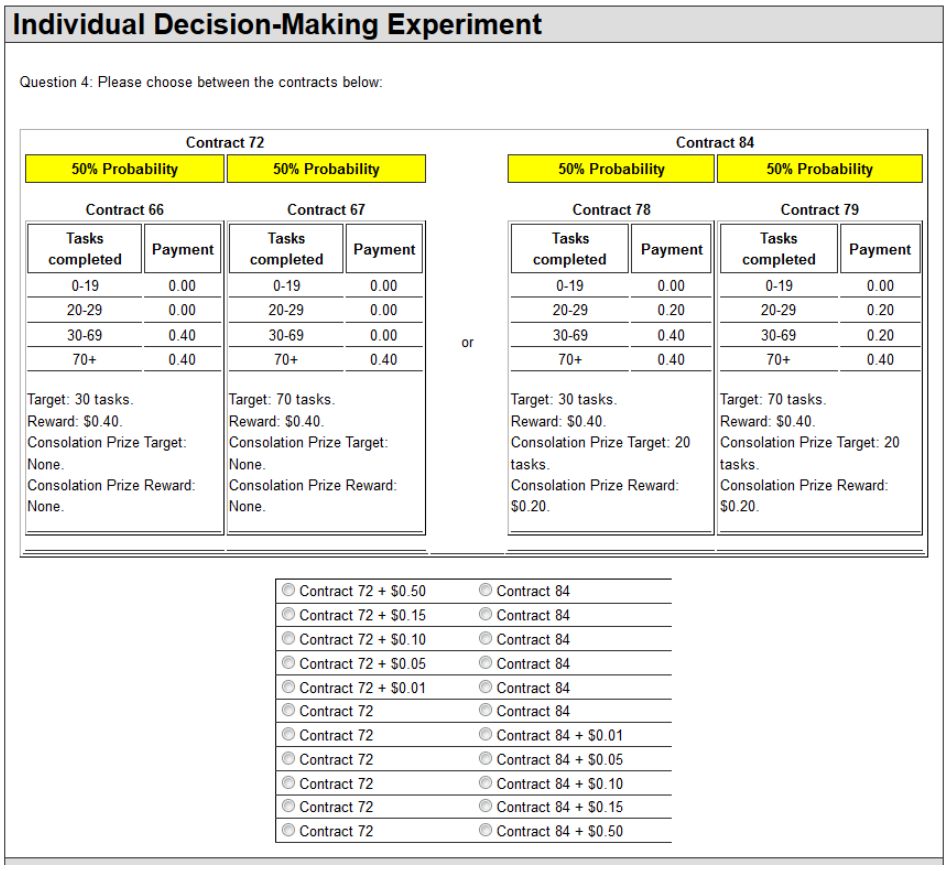


Figure 4: Sample random contracts question

An example of a question involving random contracts can be found in Figure 4. For all Low and Flex contracts in this treatment, the "low" option was the same: \$0.20 for completing 20 tasks. In High and Flex contracts, the "high" option paid \$0.40, but the number of tasks required to earn the \$0.40 changed between contracts: it was either "easy" (requiring 30 tasks), "hard" (requiring 70 tasks), or "random". If the high option was 'random', then the contract involved a lottery which assigned the subject to either the 'easy' or the 'hard' parameters for the high option with equal probability. If the subject selected a random contract, then the number of tasks necessary for the high option was determined before it was implemented.

There were two kinds of random contracts used in this treatment: "High-random" which was a 50/50 lottery between the contract with only the 'high-easy' option and the contract with only the 'high-hard' option (which we will call "H-random" or "H-r"), and "Flex-random" which was a lottery between the Flex contract formed with the low option and the 'high-easy' option and the Flex contract composed of the low option and the 'high-hard' option ("Flex-random" or "F-r"). Nonrandom contracts that utilized the "easy" option are referred to as H-e and F-e, while those that utilized the "hard" option are referred to as H-h and F-h.

For this treatment, as in the baseline, subjects completed three exogenous task sections, again in a random order. Two of these involved degenerate contracts: 'Flex-easy' and 'Flex-hard'. The remaining task section used the the 'Flex-random' contract. For this section, the subject was informed that the contract they were using in this section involved randomization, and reminded that the computer would generate a random number which would determine which of the two subcontracts they would use, either the Flex-easy or the Flex-hard. On the next screen they were shown the random number generated by the computer, and reminded which contract they would be using in the task section.

Subjects then responded to MPL questions asking for their preferences between Flex and Low contracts and between Flex and High contracts for the three different types of Flex (and High) menus. That is, they responded to MPLs comparing L/F-e, H-e/F-e, L/F-h, H-h/F-h, L/F-r, and H-r/F-r.[13] If uncertainty does lead to a preference for flexibility, we would expect to see more such behavior for the random contracts than for the easy or hard contracts.

As in the main treatment, one of these choices was actualized, and the subject

[13]The order of presentation was randomized between subjects.

completed the fourth and final task section using this contract. If the realized contract was one involving randomization, then the subject was reminded of their choice and that the computer would generate a random number to determine the contract to be used.

4 Data Overview

In total, 239 subjects participated in the main experimental treatment and 149 in the random contracts treatment, between November 9, 2013 and July 12, 2014. All subjects were recruited through Amazon's Mechanical Turk, and completed the experiment over the Internet. Subjects earned an average of $1.35, including a participation fee of $0.25.

4.1 Data Quality

While all experimental data contains a certain amount of noise, the use of Mechanical Turk makes our data particularly likely to contain uninformative responses. The prevailing norm on MT is to perform small and repetitive tasks for small sums of money, and while this prepares the subjects well for the tasks used in this experiment, it also potentially rewards a meta-strategy of not exerting too much effort on any given HIT, instead focusing on completing a larger volume of HITs in a given time. This accordingly increases the concern that the subjects were clicking through the MPL questions without serious consideration of the choices being made. Because of this, we are particularly interested in identifying subjects who were not paying attention when choosing between contracts, so their data can be excluded from further analysis.

To screen these subjects from our data, the MPL method provides two potential tools. The first, as previously mentioned, is that the right-hand side option becomes more attractive relative to the left-hand side option as one goes down the list. This means that subjects' choices should never cross from right to left. The second is that the side payments used on the first and last lines of each question are sufficiently large to overwhelm any marginal earnings from having one contract over another. The additional payments on the top and bottom lines of each question are $0.50, which is more than can be made with any given contract under consideration, and so the $0.50 payment should always dominate the contracts in question. Accordingly,

subjects should always select the option on the left on the first line, and the option on the right on the last line of the question, leading to exactly one crossover from left to right.

In the main experimental treatment we used these checks to exclude subjects based on their responses to the three MPL questions related to the contracts they experienced during the exogenous contracts section (L20/H, L20/F20, H/F20). In these questions 89% of subjects never switched from right to left, while 56% always selected the left option on the first line and the the right option in the last line. Excluding subjects that fail one of these checks left 124 subjects or 52% of the original. While this rate of exclusion is high relative to most laboratory experiments, it is unsurprising given the small stakes and noise inherent in the MT environment. Using equivalent criteria[14] we retained 61 subjects (41%) from the random contracts treatment. It should also be noted that our screening procedure in fact reduces the amount of preference for flexibility we observe: 43% of all subjects exhibited strict preference for flexibility amongst H, F20 and L20 contracts, compared to 35% of subjects who survived screening on these three questions.

4.2 Identifying Menu Preferences and Stochastic Choice

We now discuss how we use the data generated by the experiment to identify the menu preferences and stochastic choice functions of our subjects.

4.2.1 Menu Preference

It is useful to interpret the contracts in our experiment as being menus of options defined by the tuple (n, p) where n refers to number of tasks completed and p to the monetary payments received. All menus contain the option $(0, 0)$: i.e. to do zero tasks for no payment. LX and FX menus additionally contain the option $(X, 0.20)$. H and FX menus additionally contain the option $(50, 0.40)$. In this sense, the FX contract is the union of the LX and H menus.[15]

[14] Because the analysis of the random contracts treatments require answers to all six questions asked, these subjects are screened on six rather than three questions, leading to a somewhat higher exclusion rate.

[15] This is true if one considers non-dominated options. The LX menu contains the option to do the high number of tasks for the low reward (i.e. (50, \$0.20)), whereas the FX menu does not. However all options other than the three described above are strictly dominated, in the sense that the subject could exert less effort for the same reward. We will restrict our attention to the three non-dominated

We can identify preferences between the various menus using the MPL questions. We will say that Contract A is weakly preferred to Contract B if it is chosen in the case of no side payments. Moreover, we can identify strict preferences using the assumption that a subject who is prepared to pay a positive amount to have Contract A rather than Contract B must strictly prefer A to B.

Note that our MPL allows us only to identify strict preferences in which the difference in value between the two options is greater than \$0.01. Preferences that are weaker than this will be classified as indifference. Thus, our estimates provide a lower bound on strict preference for flexibility.

4.2.2 Identification of Stochastic Choice

In a given task section the menu of possible task/payment options was governed by the contract in effect in that task section. Subject choice from a menu was measured by the number of tasks they choose to complete.[16] We estimated choices from menus using data from the task sections in which the contracts are exogenously determined.

In practical terms, it was not possible to collect enough data to estimate subject-level stochastic choice functions. Completing the real-effort task is costly in terms of both time and effort, and so in our design we only observe one choice from a given contract per subject. We therefore estimated stochastic choice functions by pooling data across subjects. To do so we grouped subjects into equivalence classes based on their expressed preferences over menus. Details on how the subjects were pooled are given in Section 5.

5 Results

5.1 Preference for Flexibility

Uncertainty about the relative cost of effort in future task sections could lead our subjects to exhibit a strict preference for flexibility, as discussed in Section 2. In

options described above.

[16] Choices were measured as the highest non-dominated number of tasks completed. So, for example, in an F20 contract, subjects who completed 25 tasks were classed as choosing to complete 20. Over all subjects and all task sections, subjects completed an average of 3.6 'unnecessary' tasks per section. Restricting attention to subjects who tended not to complete extra tasks (i.e. completed more than 5 unnecessary tasks no more than once) removes 18 subjects, but does not substantially change our results.

our experiment a subject exhibited strict preference for flexibility if they strictly preferred the FX contract to both the LX contract and the H contract. The preference uncertainty model also requires set monotonicity, which implies that the FX contract must be weakly preferred to both the LX and the H contracts.

We used the data from the pairwise choices between F20, L20 and H to group subjects into five categories based on their revealed preferences:

1. Preference for Flexibility: $F20 \succ L20$ and $F20 \succ H$

2. Standard: Either $F20 \succ H$ and $L20 \succ H$ with no strict preference between $F20$ and $L20$, or $F20 \succ L20$ and $H \succ L20$ and no strict preference between $F20$ and H

3. Indifferent: No strict preference between any contracts

4. Commitment: Either $H \succ F20$ or $L20 \succ F20$, with no intransitivity

5. Intransitive: $(X \succ or \succeq Y$ and $Y \succ Z)$ or $(X \succ Y$ and $Y \succ or \succeq Z)$ but not $X \succ Z$, for some combination of menus X, Y, and Z

Appendix A provides a table which lists the categorization for all non-intransitive preference profiles.

The preferences of 'Standard' and 'Indifferent' subjects could be explained by a model of menu choice in which the DM has no preference uncertainty - i.e. by a model of the form of Equation 1 in which there is only one state (see for example Kreps (1979)). The 'Preference for Flexibility', or 'PFF' subjects can only be explained by this model if there is preference uncertainty. 'Commitment' subjects cannot be explained in the framework of Equation 1, although they could be explained by a model of temptation and self-control (for example Gul and Pesendorfer (2001)).[17] Subjects in the 'Intransitive' category cannot be explained by any utility-based model of choice between menus.

Table 1 shows the breakdown of subjects across these five categories. It shows that a strict preference for flexibility was common in our task, with 35% of subjects falling into this category. Overall, the large majority of subjects (85%) behaved in line with the preference for flexibility model of Equation 1. Of those that did not,

[17] Another class of models that could capture these subjects is that which includes models of regret, for example Sarver (2008).

Table 1: *Breakdown of subject type in baseline specification*

Type	N	Percent	Benchmark I	p-value	Benchmark II	p-value
Flexibility	43	35%	17%	0.00	6%	0.00
Standard	40	32%	17%	0.00	6%	0.00
Indifferent	23	19%	25%	0.12	13%	0.06
Commitment	7	6%	42%	0.00	16%	0.00
Intransitive	11	9%	-	-	59%	0.00

9% violated the transitivity condition. Only 6% of subjects fall into the commitment category, violating the set monotonicity condition.

We benchmarked our results in two ways, both in the spirit of Bronars (1987) and Beatty and Crawford (2011). Benchmark I in Table 1 is the fraction of subjects that would fall into each category if preference profiles were drawn (uniformly) at random from the possible transitive preference profiles in our experiment. Benchmark II is the distribution of preference types that would result from subjects randomizing between the four possible preferences $(\succ, \succeq, \preceq, \prec)$ in each set of pairwise choices. In both cases, the p-values report the result of the binomial test that the observed frequency is equal to the benchmark frequency. In both cases the fraction of subjects who exhibited preference for flexibility is significantly higher than the benchmark, while the fraction of subject who exhibited preference for commitment is significantly lower.

It is notable how little preference for commitment we observe in our subjects. Following Gul and Pesendorfer (2001), much of the theoretical work on menu preference has been related to issues of temptation and self-control, and how these can generate a strict preference for smaller choice sets. Indeed, Kaur, Kremer and Mullainathan (2010) demonstrates a preference for commitment in a task similar to ours in a study of data entry personnel in India. As we discuss below, our work suggests that preference uncertainty may act as a powerful offsetting force to preference for commitment.

Preference for flexibility is also present for other values of the low threshold. For the F15/L15 and F10/L10 contracts, rates are similar to that of the F20/L20 case described above, with 34% and 35% of well-behaved subjects exhibiting strict preference for flexibility. At the F5/L5 level the rate is somewhat lower at 30%. Overall, 48% of well-behaved subjects exhibited preference for flexibility at some value of the low threshold.

Note that the preference uncertainty model does not provide any prediction as to whether preference for flexibility should increase or decrease as the value of the low threshold changes. On one hand, as the low threshold gets more appealing, subjects who previously had a preference for flexibility might switch to unequivocally preferring the low threshold. On the other, subjects who previously unequivocally preferred the high contract might switch to preferring flexibility. However, willingness to pay to include the Low element in the menu should be (weakly) increasing as the number of tasks required to complete it falls. This is the case in our data: the average willingness to pay to add L is \$0.039, \$0.050, \$0.051 and \$0.054 for 20, 15, 10, and 5 tasks, respectively.[18] On an individual level, 69% of subjects exhibited a weakly increasing willingness to pay to add the low contract.

5.2 Stochastic Choice

We next examine whether the stochastic choice data generated by our subjects is consistent with the preference uncertainty model. This first requires us to ensure that the data is consistent with some underlying preference uncertainty, then second to ensure that stochastic choice and menu preferences can be rationalized by the same underlying preference uncertainty.

Because we do not observe enough choices from the same menu to meaningfully identify individual-level stochastic choice functions, we instead constructed equivalence classes of subjects based on their responses to the menu preference questions. We make use of three different partitions on our set of subjects in the following analysis. First, we grouped subjects into the PFF, Standard and Indifferent categories defined above (we do not have enough subjects in the Commitment and Intransitive categories to construct meaningful stochastic choice functions). Next we divided subjects into those who strictly preferred F20 to H and those who did not. Finally we grouped subjects into those that strictly preferred F20 to L20 and those who did not.

[18]The difference between L20 - the contract used in the exogenous contracts section - and the other L's is statistically significant (Wilcoxon signed rank p-values of p=0.01, 0.02 and 0.001 respectively). The difference in willingness to pay between the other LX contracts moves in the anticipated direction, but does not attain statistical significance.

Table 2: *Frequency of choice by contract*

Number of subjects who choose to do:			
	20 tasks in L20	20 tasks in F20	
All (N=124)	114	30	p= 0.00
PFF (N=43)	39	12	p= 0.00
Standard (N=40)	38	10	p= 0.00
Indiff (N=23)	21	4	p= 0.00
	50 tasks in H	50 tasks in F20	
All (N=124)	79	77	p= 0.83
PFF (N=43)	31	28	p= 0.58
Standard (N=40)	27	29	p= 0.63
Indiff (N=23)	11	11	p= 1.00
	0 tasks in H	0 tasks in F20	
All (N=124)	45	17	p= 0.00
PFF (N=43)	12	3	p= 0.04
Standard (N=40)	13	1	p= 0.00
Indiff (N=23)	12	8	p= .22
	0 tasks in L20	0 tasks in F20	
All (N=124)	10	17	p= 0.14
PFF (N=43)	4	3	p= 1.00
Standard (N=40)	2	1	p= 1.00
Indiff (N=23)	2	8	p= 0.07

5.2.1 Choice Monotonicity

As discussed in Section 2 the first key test is choice monotonicity: adding additional options to a set can only weakly diminish the probability of choosing the options previously in that set. For this experiment, both the L20 and H menus are subsets of the F20 menu, so the probability of choosing to do 0 tasks in the F20 menu must be lower than the probability of doing so in either the L20 or the H menus, and the probability of doing 20 (50) tasks must also be less in the F20 menu than it is in the L20 (H) menu. Table 2 displays the relevant frequencies with which subjects chose elements from those menus for all subjects, as well as the PFF, Standard and Indifferent subsets. Reported p-values reflect the sign test.

Table 2 shows that subject behavior was broadly in line with choice monotonicity. For almost all groups and and options, the probability of choosing that option is either significantly lower in the larger menu (for example, for all subjects the probability of

Table 3: *Proportions of subjects who pay to add an element to a menu who choose that element in the exogenously imposed F20 contract*

Subjects who:	Fraction doing 20 tasks in F20	N	p-value
Do not strictly prefer F20 to H	0.09	57	p=0.00
Strictly prefer F20 to H	0.37	67	

Subjects who:	Fraction doing 50 tasks in F20	N	p-value
Do not strictly prefer F20 to L20	0.42	53	p=0.00
Strictly prefer F20 to L20	0.77	71	

doing 20 tasks was significantly lower in the F20 contract than in the L20 contract) or there is no significant difference (for example, more Standard subjects performed 50 tasks in the F20 menu than in the H menu (29 vs 27) but the related sign test has a significance level of 0.625).

The exception is that there is weak evidence that Indifferent subjects chose to do 0 tasks more often in the F20 contract than in the L20 contracts. This could be the result of a tie-breaking rule in the face of indifference which favors 0 tasks in the larger choice set. Moreover, the number of subjects who exhibit such behavior is small, and the difference is significant only at the 7% level.

5.2.2 Consequentialism

Consequentialism requires that DMs only strictly prefer a larger choice set if the additional options are at least sometimes chosen from the resulting menu. In our experiment, this means that if a subject strictly preferred the F20 menu to either the L20 or the H menus, then they must exercise the additional option this grants them at least some of the time.

Because this prediction involves the stochastic choice function, we must utilize the equivalence classes in order to test it. The top panel of Table 3 divides subjects into those who exhibited a strict preference for F20 over H and those that did not. For each group it shows the fraction of subjects who did 20 questions when exogenously given the F20 contract. The second panel divides subjects into those who exhibited a strict preference for F20 over L20 and those that did not, reporting the fraction of subjects who performed 50 tasks in the exogenously provided F20 contract.

It is clear from Table 3 that the (rather moderate) requirement of the Consequentialism condition is satisfied. Subjects who strictly preferred F20 to H did 20 tasks

Table 4: *Breakdown of subject choices in exogenous F20 contract*

	PFF		L20/Indiff		H/Indiff	
	F20≻L20, F20≻H		Not F20≻L20		Not F20≻H	
N=	44 subjects		40 subjects		46 subjects	
Completing at least	Number	Percent	Number	Percent	Number	Percent
0 tasks	3	.07	8	.20	9	.20
20 tasks	12	.28	14	.35	4	.09
50 tasks	28	.65	18	.45	33	.72

37% of the time in the F20 contract, while those that that strictly preferred F20 to L20 did 50 tasks 77% of the time in the F20 contract.

A more stringent condition is that subjects who are prepared to pay to have an option included in their choice set should use it more than those that are not. While this is not a requirement of the model characterized by AS, it is implied if the number of subjects who are indifferent between the different elements of the Flex contract is small.[19] Table 3 shows that subjects who paid to add an element to a menu were significantly more likely to exercise that option, with a Mann-Whitney p-value that is significant at the 0.01 level.

A corollary to the Consequentialism condition is that subjects who exhibit preference for flexibility should make use of both the 20 and 50 question levels. Table 4 shows that this is the case. The first two columns show the distribution of choices made in the F20 contract by PFF subjects. The second two columns show the same distribution for Standard/Indifferent subjects who did not have a strict preference for F20 over L20. The last two show the distribution for such subjects who did not have a strict preference for F20 over H. Subjects who exhibited preference for flexibility made significant use of both the 20 and 50 task levels (28% and 65% respectively). Moreover, they made use of the 50 task level more than subjects who showed no strict preference for the F20 over the L20 contract (p=0.067) and made use of the 20 task level more that those who showed no strict preference for F20 over the H contract (p=0.026).

[19]The number of indifferent subjects is an upper bound on the number of subjects who do not strictly prefer F20 over H (L20) but complete 20 (50) tasks. For a further discussion see Section 5.2.3.

Table 5: *Frequency of Preferring to add L20 (H), for subjects that do L20 (H) in F*

	Menu Preference:	All Subjects	Non-Indifferent
Subj. that do 20 tasks in F20	$F20 \succ H$	0.83	0.96
	$F20 \succ L20$	0.40	0.46
Subj. that do 50 tasks in F20	$F20 \succ H$	0.47	0.55
	$F20 \succ L20$	0.71	0.83

5.2.3 Responsive Menu Preferences

The converse of the Consequentialism condition would be that $\lambda^{A\cup\{x\}}(x) > 0$ implies $A \cup \{x\} \succ A$. In our experiment, this would imply that subjects who chose to do 20 tasks when the F20 contract was exogenously given should strictly prefer F20 over H, while subjects who did 50 tasks should strictly prefer F20 to L20. Table 5 shows the menu preferences of these two groups of subjects. It shows that a large majority (83%) of subjects who did 20 tasks did indeed strictly prefer F20 to H (relative to 47% of subjects who did 50 tasks - significantly different at the 0.1% level, p=0.001). Similarly, 71% of subjects who did 50 tasks preferred F20 to L20 (relative to 40% of the subjects who did 20 tasks - significantly different at the 1% level, p=0.003).

As AS point out this condition is overly restrictive: it could be the case that the subject was indifferent between (for example), performing 20 tasks for \$0.20 or performing 50 tasks for \$0.40, but preferred both to performing 0 tasks for \$0.00. Such a subject would choose to do one of 20 or 50 tasks, but would not pay for that effort level to be included in their contract. Thus, the fact that some subjects who did 20 tasks (50 tasks) were not prepared to pay for F20 over H (F20 over L20) may be due to indifference, or a strength of preference of less than \$0.01, the smallest increment measurable in our multiple price lists. Such subjects should display no strict preferences between contracts - i.e. they should fall into the 'Indifferent' category. The final column of Table 5 shows that dropping these subjects increases compliance with the Responsive Menu Preference condition - to 96% for subjects who did 20 tasks from the F20 contract and to 83% for subjects who did 50 tasks.

5.3 Exogenous Uncertainty

Our final experiment tested whether exogenously introducing uncertainty about the nature of the contract increases preference for flexibility. If a subject expresses pref-

erence for flexibility in the Easy or the Hard contracts in the 'Random Contracts' treatment, then they should also have a preference for flexibility in the Random contract. Furthermore, there may be subjects who exhibit preference for flexibility for the Random, but not for either the Easy or Hard contracts. This may include subjects who have no intrinsic preference uncertainty - for example a subject who always prefers to do 30 tasks for \$0.40 over 20 tasks for \$0.20, but always prefers to do 20 tasks for \$0.20 over 70 tasks for \$0.40 should exhibit preference for flexibility only for the Random contract. This means the fraction of subjects expressing preference for flexibility should be (weakly) higher in the Random contracts than in either the Easy or the Hard contracts: the introduction of the exogenous uncertainty over the difficulty of the 'hard' level should induce a higher fraction of subjects to express a preference for flexibility.

Of the 61 well-behaved subjects in this experiment, 41% exhibited preference for flexibility for the Random contract, significantly higher than for the Easy (23%) and Hard (16%) contracts (McNemar p-values of 0.04, 0.01). At the individual level, 84% of subjects behaved in line with the prediction that preference for flexibility in either the Easy or Hard contracts should lead to preference for flexibility in the Random contract. 28% of subjects exhibited preference for flexibility only for the Random contract.

6 What Is the Source of Subject Uncertainty?

In this experiment we find a considerable amount of preference for flexibility, despite the fact that there was a relatively small temporal gap between the choice of contract and its implementation. This begs the question: were subjects really uncertain about relative effort costs, or did subjects demand the flexible contract because they feared they would be unable to complete level H? It is potentially the case that subjects demanded the FX contract because they were uncertain not about their future preferences, but about their future *ability* to complete the tasks.

Although we do not know what the subjects were thinking at the time they made their choices, we can test whether such a fear about ability to finish is well-founded. That is, did subjects seem to have had a hard time completing 50 tasks within the time limit when they attempted to do so, meaning that they may sensibly prefer to hedge against the possibility that they try and fail to do so in the future?

We argue that this is not the case. Of the 248 exogenous tasks sections in which subjects could have potentially completed level H (in the sections completed before choosing between contracts), there were only 13 instances (5.2%) in which a subject used more than 85% of their 15 minute time limit without successfully completing level H. It is also possible that subjects realized that they would have been unable to complete level H at their current rate, and so quit the section before spending too much time on it. However, this also seems to have happened infrequently, as there were only 29 instances (11.7%) in which subjects would not have completed at least 50 tasks at their current rate.[20] Of course, this is an upper bound on the number of times a subject was attempting to complete the higher number of tasks but would have been unable to do so, as the task speed is endogenous; anyone who intended to complete only the lower number of tasks could feel free to take as long as they wish.

It could still be the case that the subjects who expressed a preference for flexibility were those relatively few subjects who perhaps would have liked to complete 50 tasks but would have been unable to do so. This is not the case. Subjects who experienced a task section in which they would not have completed 50 tasks at their current rate expressed lower preference for flexibility than subjects who did not (0.17 vs 0.39, Fisher exact p=0.087).[21]

7 Related Literature

This work is positioned at the intersection of two theoretical areas: the literature on preference for flexibility and the literature on random choice.

The models investigated in this project are Dekel et al. (2007), Gul and Pesendorfer (2006), and Ahn and Sarver (2013). DLRS extends the previous flexibility model developed in Kreps (1979), and has become the workhorse model of preference for flexibility. In DLRS a desire for larger choice sets at the time of choice from menus is driven by uncertainty over prevailing preferences at the time of choice from menus. We focus on the version of DLRS which entails flexibility-seeking (as in AS). GP models stochastic choice from menus as the result of DMs maximizing state-dependent utility functions given a stochastic underlying state space that is unobserved by the

[20] Computed for subjects who completed at least five tasks.

[21] Similarly, those subjects who at some point used >85% of their time exhibited less preference for flexibility, although the difference is not statistically significant (0.27 vs 0.35, Fisher exact p=.75).

researcher. AS provides the axiomatic foundation formally joining these two models into a unified theory in which the same underlying (unobserved) state space that engenders preference for flexibility in menu choice also drives stochastic choice from menus. Thus, AS nests GP and DLRS, and accordingly is the key model used here; the aspects of each these papers most relevant to this project are discussed in greater detail in Section 2.

There is a relatively small but growing literature on menu preference in laboratory and field experiments. In the laboratory, Sonsino and Mandelbaum (2001) document a preference for increased menu size and experimentally examine the tradeoff between a desire for larger choice sets and potential aversion to decision complexity by eliciting subjects' values for stochastic asset portfolios that vary in the number of options they contain. In their across-subjects design, they find that subjects placed a higher value on larger menus compared to menus that are strict subsets, which they interpret as a form of flexibility-seeking, based on the nature of the elements added. This differs from the definition of flexibility we consider, which requires a DM to value the union of two menus strictly higher than either of the submenus.

The related (though in some sense opposite) phenomenon of preference for smaller menus has also been experimentally documented in several environments, ranging from field experiments involving savings plans (Ashraf, Karlan and Yin (2006) and Giné, Karlan and Zinman (2010)) to laboratory experiments with real-effort tasks (Houser et al. (2010) and Augenblick, Niederle and Sprenger (2013)). Most closely related to our experiment is Kaur, Kremer and Mullainathan (2010), which studies commitment-seeking behavior in a field experiment involving piece-rate data-entry workers in India who similarly choose between different contract structures.

Stochastic choice has been experimentally documented at least as far back as Tversky (1969). Numerous experiments, often intended to test Expected Utility and relaxations thereof, have found that subjects often change their answer to a question when it is repeated, and have treated this data as a form of 'reliability check' as in Camerer (1989), or used this data to estimate a model with 'mistakes' or white noise variation, as in Hey and Orme (1994). Later work has treated the stochastic choice as an explicit object of study, as in Hey (2001) which investigates whether the choices have reduced variability with increased repetition. More recently, Regenwetter and Davis-Stober (2012) tests whether it is possible for seemingly intransitive choice data to come from a collection of underlying weak orders (as in GP), and finds that

the data in their sample is consistent with that possibility. On the other hand, Agranov and Ortoleva (2013) also investigates stochastic choice behavior, but finds that subjects actually paid to use a costly randomization device, which they argue is a type of stochastic choice that is not explained using random utility, but fits better with models of 'hedging'.

8 Conclusion

The objective of this project was to gather evidence of preference for flexibility in a controlled setting, determine whether preference uncertainty could offer a unified description of choice between and from menus, and provide evidence of a causal role for preference uncertainty in determining menu preferences. We argue that the experimental data we gather from a quasi-field setting achieves all three of these aims: 48% of our subjects exhibited strict preference for flexibility, subject behavior fits well with the unified AS model, and introducing exogenous uncertainty increases preference for flexibility.

Finding substantial preference uncertainty suggests that it may be important to consider its effect when designing employment contracts. One particularly interesting possibility is that this preference uncertainty may work against demand for commitment driven by temporally inconsistent preferences, for example of the type discussed by Gul and Pesendorfer (2001). The presence of both preference uncertainty and time inconsistencies can generate a tension for the design of optimal contracts, as discussed in Amador, Werning and Angeletos (2006): on one hand, DMs want to restrict their future choices to reduce self-control problems. On the other hand, they want to leave themselves the flexibility to respond to preference shocks. This indicates that are potential welfare gains to be made by incorporating a degree of flexibility into commitment devices, and indeed there already exist services which offer more flexible forms of commitment.[22] Our results suggest that contracts that consider both effects may offer significant welfare gains over contracts which do not take preference uncertainty into account.

[22]For example, https://www.beeminder.com/.

References

Agranov, Marina and Pietro Ortoleva. 2013. "Stochastic Choice and Hedging*.".

Ahn, David and Todd Sarver. 2013. "Preference for Flexibility and Random Choice." *Econometrica* 81(1):341–361.

Amador, Manuel, Ivan Werning and George-Marios Angeletos. 2006. "Commitment vs. flexibility." *Econometrica* 74(2):365–396.

Ashraf, Nava, Dean Karlan and Wesley Yin. 2006. "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines." *The Quarterly Journal of Economics* (May):635–672.

Augenblick, Ned, Muriel Niederle and Charles Sprenger. 2013. "Working Over Time : Dynamic Inconsistency in Real Effort Tasks*.".

Beatty, Timothy K M and Ian A Crawford. 2011. "How Demanding Is the Revealed Preference Approach to Demand ?" *American Economic Review* 101(October):2782–2795.

Bronars, Stephen. 1987. "The Power of Nonparametric Tests of Preference Maximization." *Econometrica* 55(3):693–698.

Camerer, Colin F. 1989. "An Experimental Test of Several Generalized Utility Theories." *Journal of Risk and Uncertainty* 2(1):61–104.

Caplin, Andrew and Mark Dean. 2014. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." (January):1–40.

Dekel, E, BL Lipman and Aldo Rustichini. 2001. "Representing preferences with a unique subjective state space." *Econometrica* 69(4):891–934.

Dekel, Eddie, BL Lipman, Aldo Rustichini and Todd Sarver. 2007. "Representing Preferences with a Unique Subjective State Space: A Corrigendum1." *Econometrica* 75(2):591–600.

Giné, Xavier, Dean Karlan and Jonathan Zinman. 2010. "Put your money where your butt is: a commitment contract for smoking cessation." *American Economic Journal: Applied ...* 2(4):213–235.

Goodman, Joseph, Cynthia Cryder and Amar Cheema. 2013. "Data Collection in a Flat World : The Strengths and Weaknesses of Mechanical." *Journal of Behavioral Decision Making* 26:213–224.

Gul, F and W Pesendorfer. 2001. "Temptation and self-control." *Econometrica* 69(6):1403–1435.

Gul, F and W Pesendorfer. 2006. "Random expected utility." *Econometrica* .

Hey, John and Chris Orme. 1994. "Investigating Generalizations of Expected Utility Theory Using Experimental Data." *Econometrica* 62(6):1291–1326.

Hey, John D. 2001. "Does Repetition Improve Consistency ?" *Experimental Economics* 4:5–54.

Houser, Daniel, Daniel Schunk, Joachim Winter and Erte Xiao. 2010. "Temptation and Commitment in the Laboratory*." *University of Zurich Working Paper* (488).

Kaur, S, M Kremer and S Mullainathan. 2010. "Self-Control at Work: Evidence from a Field Experiment.".

Kreps, DM. 1979. "A representation theorem for" preference for flexibility"." *Econometrica: Journal of the Econometric Society* 47(3):565–577.

Mason, Winter and Siddharth Suri. 2012. "Conducting behavioral research on Amazon's Mechanical Turk." *Behavior Research Methods* 44(1):1–23.

Paolacci, Gabriele, Jesse Chandler and Leonard N Stern. 2010. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making* 5(5):411–419.

Regenwetter, Michel and Clintin Davis-Stober. 2012. "Behavioral Variability of Choices Versus Structural Inconsistency of Preferences." *Psychological Review* 119(2):408–416.

Sarver, Todd. 2008. "Anticipating Regret: Why Fewer Options May Be Better." *Econometrica* 76(2):263–305.

Sonsino, Doron and Marvin Mandelbaum. 2001. "On Preference for Flexibility and Complexity Aversion: Experimental Evidence." *Theory and Decision* 51:197–216.

Tversky, Amos. 1969. "Intransitivity of Preferences." *Psychological Review* 76(1):31–48.

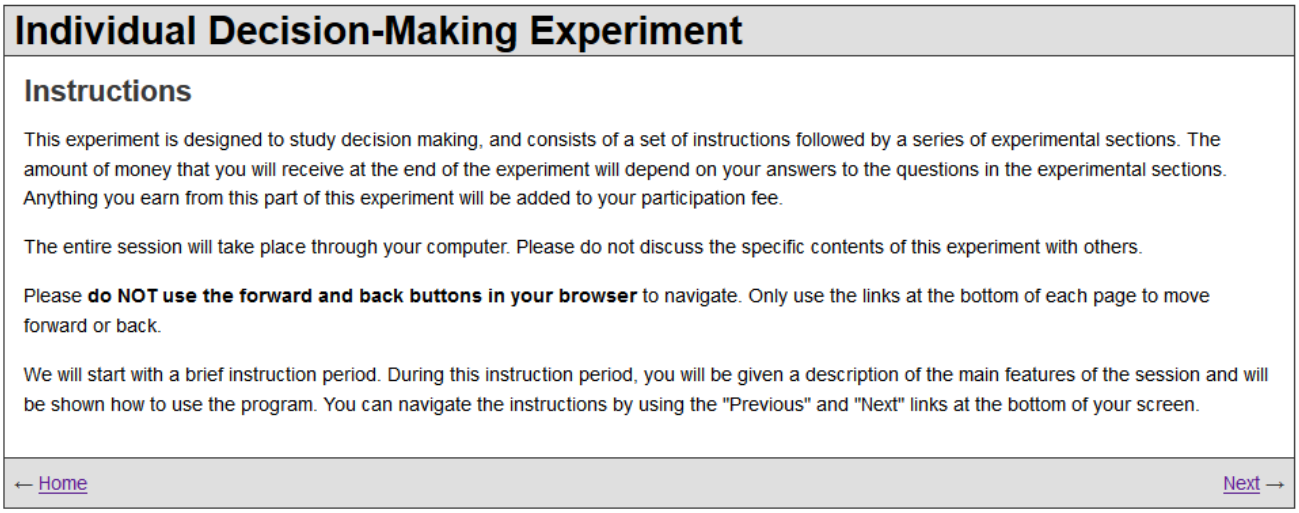
A Preference Profiles

Type of Preference	Allowable Preferences										
Standard-L	F	$\succ$	H		F	$\succeq$	L		H	$\prec$	L
	F	$\succ$	H		F	$\preceq$	L		H	$\prec$	L
Standard-H	F	$\succeq$	H		F	$\succ$	L		H	$\succ$	L
	F	$\preceq$	H		F	$\succ$	L		H	$\succ$	L
PFF	F	$\succ$	H		F	$\succ$	L		H	$\preceq$	L
	F	$\succ$	H		F	$\succ$	L		H	$\prec$	L
	F	$\succ$	H		F	$\succ$	L		H	$\succeq$	L
	F	$\succ$	H		F	$\succ$	L		H	$\succ$	L
Commitment-L	F	$\succeq$	H		F	$\prec$	L		H	$\prec$	L
	F	$\succ$	H		F	$\prec$	L		H	$\prec$	L
	F	$\preceq$	H		F	$\prec$	L		H	$\prec$	L
Commitment-H	F	$\prec$	H		F	$\succeq$	L		H	$\succ$	L
	F	$\prec$	H		F	$\succ$	L		H	$\succ$	L
	F	$\prec$	H		F	$\preceq$	L		H	$\succ$	L
Commitment-Either	F	$\prec$	H		F	$\prec$	L		H	$\preceq$	L
	F	$\prec$	H		F	$\prec$	L		H	$\prec$	L
	F	$\prec$	H		F	$\prec$	L		H	$\preceq$	L
	F	$\prec$	H		F	$\prec$	L		H	$\succ$	L
Indifferent	F	$\succeq$	H		F	$\succeq$	L		H	$\succeq$	L
	F	$\succeq$	H		F	$\succeq$	L		H	$\preceq$	L
	F	$\succeq$	H		F	$\preceq$	L		H	$\succeq$	L
	F	$\succeq$	H		F	$\preceq$	L		H	$\preceq$	L
	F	$\preceq$	H		F	$\succeq$	L		H	$\succeq$	L
	F	$\succeq$	H		F	$\preceq$	L		H	$\preceq$	L
	F	$\preceq$	H		F	$\succeq$	L		H	$\succeq$	L
	F	$\preceq$	H		F	$\preceq$	L		H	$\succeq$	L
	F	$\preceq$	H		F	$\preceq$	L		H	$\preceq$	L

Appendix B: Experimental Instructions

This is an example of the instructions seen by a subject in this experiment. The instructions shown are for the random contracts treatment, as the instructions shown to subjects in this treatment are a superset of those used in the main treatment. The sections of the instructions that would not be used in the main treatment are marked.

The experiment begins with an instructional section.



Individual Decision-Making Experiment

Instructions

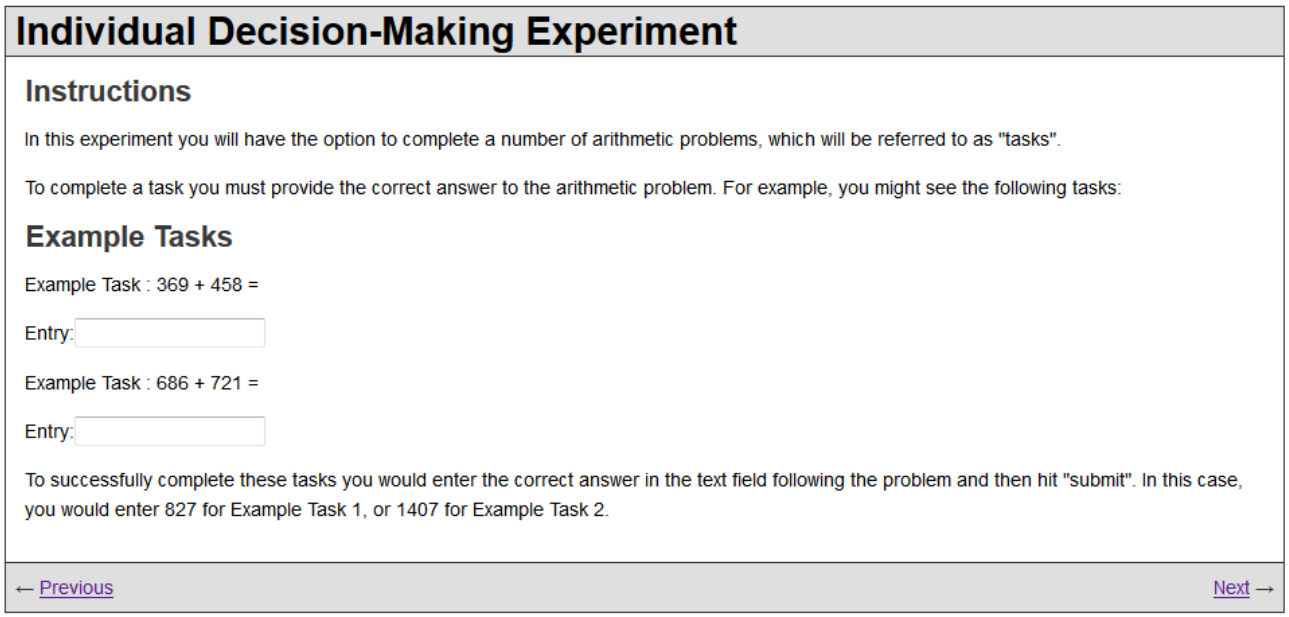
This experiment is designed to study decision making, and consists of a set of instructions followed by a series of experimental sections. The amount of money that you will receive at the end of the experiment will depend on your answers to the questions in the experimental sections. Anything you earn from this part of this experiment will be added to your participation fee.

The entire session will take place through your computer. Please do not discuss the specific contents of this experiment with others.

Please **do NOT use the forward and back buttons in your browser** to navigate. Only use the links at the bottom of each page to move forward or back.

We will start with a brief instruction period. During this instruction period, you will be given a description of the main features of the session and will be shown how to use the program. You can navigate the instructions by using the "Previous" and "Next" links at the bottom of your screen.

← Home Next →



Individual Decision-Making Experiment

Instructions

In this experiment you will have the option to complete a number of arithmetic problems, which will be referred to as "tasks".

To complete a task you must provide the correct answer to the arithmetic problem. For example, you might see the following tasks:

Example Tasks

Example Task : 369 + 458 =

Entry:

Example Task : 686 + 721 =

Entry:

To successfully complete these tasks you would enter the correct answer in the text field following the problem and then hit "submit". In this case, you would enter 827 for Example Task 1, or 1407 for Example Task 2.

← Previous Next →

Individual Decision-Making Experiment

Instructions

The experiment is organized into several sections. In each section you may complete as many or as few of the tasks as you choose. The number of tasks you complete in a section will not change your future sections in any way.

When you are completing tasks, you will notice that below the task you will see a countdown like the one below:

"Please enter a solution within 19:33."

For each experimental section, you will have **20 minutes** to complete the tasks in that section.

If you run out of time in a given section, that section will end, and you will be redirected to the next experimental section.

If you attempt to complete a task but enter an incorrect solution, that task **will not** count towards the total number of tasks you have completed successfully. If that happens, you will be alerted that the solution you provided was incorrect, and will be given a **new** task.

← Previous Next →

Individual Decision-Making Experiment

Instructions

During each experimental section you may complete as many or as few tasks as you choose.

When you have decided to stop completing tasks in a given section, click on "Finish with this section" on the lower left-hand side of the screen. Please note that you can **not** return to a section once you have completed it.

In short, each section may conclude in one of two ways. Either

- you decide to stop completing tasks in that section and click "Finish with this section", or
- you run out of time for a given section, and the section ends automatically

← Previous Next →

Instructions

In each section you will have a target number of tasks to complete. The amount that you will earn in each section is determined by a "contract" that specifies your target number of tasks, and the amount of bonus that you will be paid for that section if you meet your target, as in the example below.

Example Contract

Your target for this contract is 50 tasks. If you complete at least 50 tasks you will be paid an additional $0.40 for this section.

Contract 97

Tasks completed	Payment
0-19	0.00
20-29	0.00
30-49	0.00
50-69	0.40
70+	0.40

Target: 50 tasks.
Reward: $0.40.
Consolation Prize Target: None.
Consolation Prize Reward: None.

Some contracts also include a consolation prize if you do not reach the full target level, but still complete a certain number of tasks, as in the example below.

Example Contract

Your target for this contract is 50 tasks. If you complete at least 50 tasks you will be paid an additional $0.40 for this section.
For this contract, if you fail to meet your target but complete at least 20 tasks you will be awarded a consolation prize of $0.20 for this section.

Contract 98

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.40
70+	0.40

Target: 50 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Please note that the payment amount refers to bonus you will be paid for reaching the target level, and **NOT** the amount *per* task.

Regardless of which contract you have, you may always complete as many or as few tasks as you wish, in any section. Each section is separate, and you cannot return back to sections after starting the next section.

When you have completed the experiment, you will be given a unique code to enter into Mechanical Turk to complete the HIT. This code will be used to make sure you are paid the correct bonus amount, which will correspond to your earnings from completing tasks during the experiment. Please allow up to 48 hours for your bonus payment to be processed.

Subjects in the random contracts treatment also saw:

Sometimes you will be randomly assigned to one of two contracts. When that is the case, the computer will generate a random number between 1 and 100. If that number is between 1 and 50, you will be assigned to the first contract. If that number is between 51 and 100, you will be assigned to the second contract.

For example, you might see:

Contract 92

50% Probability	50% Probability

Contract 90

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.40
50-69	0.40
70+	0.40

Target: 30 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Contract 91

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.20
70+	0.40

Target: 70 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

This means that with a 50% probability you will use the contract on the left (Contract 90), and with a 50% probability you would use the contract on the right (Contract 91).

If you are randomly assigned to a contract, you will be told the random draw and which contract you will use **BEFORE** you begin work on that task section.

To make sure that you understand the instructions so far, we are going to ask you a few questions about contract payments. You can click "Next" to begin, or "Previous" to review earlier instructions.

Subjects then responded to a short comprehension check:

Comprehension Check

Please consider the contract shown below, and answer the following questions.

Your target for this contract is 50 tasks. If you complete at least 50 tasks you will be paid an additional $0.40 for this section.
For this contract, if you fail to meet your target but complete at least 20 tasks you will be awarded a consolation prize of $0.20 for this section.

Contract 98

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.40
70+	0.40

Target: 50 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Under this contract, if you complete **0** tasks, how much will you be paid for this section? (In cents)
Under this contract, if you complete **20** tasks, how much will you be paid for this section? (In cents)
Under this contract, if you complete **40** tasks, how much will you be paid for this section? (In cents)
Under this contract, if you complete **50** tasks, how much will you be paid for this section? (In cents)
Under this contract, if you complete **100** tasks, how much will you be paid for this section? (In cents)

If the subject answered one or more questions incorrectly, he or she was informed that they were incorrect and instructed to try again, or review the previous instructions before trying again. Upon correct completion of the quiz, they were directed to the main section of the experiment.

The main part of the experiment begins by requiring that the subject complete four practice questions, to ensure familiarity with the task before beginning the main experimental task sections.

Practice Section

In the first section, you will be asked to complete a series of practice tasks. These practice tasks are examples of the tasks that you will have the option to complete later in the experiment.

The purpose of these practice tasks is to familiarize you with the experience of completing a task.

There are 4 questions in this section.

Upon completion of the four practice questions (similar to the example tasks shown in the main text), subjects began with three experimental task sections. The ordering of these tasks sections was random across individuals; examples of section instructions are included below.

An example of the instructions in a nonrandom contract section:

Section 1

Now that you have completed the practice tasks, you will begin the first regular task section.

In this section you will be paid according to the following contract:

Your target for this contract is 70 tasks. If you complete at least 70 tasks you will be paid an additional $0.40 for this section.
For this contract, if you fail to meet your target but complete at least 20 tasks you will be awarded a consolation prize of $0.20 for this section.

Contract 91

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.20
70+	0.40

Target: 70 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Click here to Continue.

Instructions for the random contracts:

You have completed 2 of the task sections. You are about to begin task section 3.

In this section you will be paid according to the following contract:

Contract 92

50% Probability	50% Probability

Contract 90

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.40
50-69	0.40
70+	0.40

Target: 30 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Contract 91

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.20
70+	0.40

Target: 70 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Please note that this contract is different from the contract in the previous sections.
Recall from the instructions that this means that the computer will choose a number at random from between 1 and 100, and if that number is 50 or below, you will be paid according to the Contract 90. If that number is 51 or above, you will be paid according to the Contract 91.
Click "Continue" to determine your random number and contract for this section.

After clicking “Continue”, subjects saw the realization of the random number, and the contract this entailed, as in the example below:

You are using Contract 92, which is a 50% chance of using Contract 90 and a 50% chance of using 91.

Your random number is 13.

This means that in the following section you will use Contract 90.

Contract 90

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.40
50-69	0.40
70+	0.40

Target: 30 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Click here to Continue.

Following the completion of three tasks sections, subjects are given additional instructions regarding the choices the make between contracts:

In this section, like the sections you have already completed, you will have the opportunity to complete a number of tasks for payment.

However, in this section, you will be asked to **choose** between contracts in a series of questions like the one shown below.

Example Question 1: Choose between Contract #86 and Contract #74:

Target: 20 tasks.	Target: 40 tasks.
Reward: $0.20.	Reward: $0.40.
Consolation Prize Target: None.	Consolation Prize Target: 20 tasks.
Consolation Prize Reward: None.	Consolation Prize Reward: $0.20.

Contract 86		Contract 74	
Tasks completed	**Payment**	**Tasks completed**	**Payment**
0-19	0.00	0-19	0.00
20-39	0.20	20-39	0.20
40-59	0.20	40-59	0.40
60+	0.20	60+	0.40
Target: 20 tasks. Reward: $0.20. Consolation Prize Target: None. Consolation Prize Reward: None.		Target: 40 tasks. Reward: $0.40. Consolation Prize Target: 20 tasks. Consolation Prize Reward: $0.20.	

○ Contract 86 + $0.50	○ Contract 74
○ Contract 86 + $0.15	○ Contract 74
○ Contract 86 + $0.10	○ Contract 74
○ Contract 86 + $0.05	○ Contract 74
○ Contract 86 + $0.01	○ Contract 74
○ Contract 86	○ Contract 74
○ Contract 86	○ Contract 74 + $0.01
○ Contract 86	○ Contract 74 + $0.05
○ Contract 86	○ Contract 74 + $0.10
○ Contract 86	○ Contract 74 + $0.15
○ Contract 86	○ Contract 74 + $0.50

Each line asks you to choose between two contracts, possibly with some additional payment. Please note that in each line of the questions, the contracts are the **same** - only the additional payments change from line to line.

Each of these lines acts as a separate question, so you should consider each line individually, and make a choice for **EACH** line.

After you have submitted your choices for all of the questions in this section, the computer will select **ONE** of those questions at random, and then **ONE** of the lines from that question at random, and your choice on that line will determine your contract for this section of the experiment. For that reason you should consider each line of the question, as all lines are equally likely to be chosen as the line that determines your contract.

For example, suppose the top line of this question were the one randomly selected to determine your contract, and suppose that you had selected "Contract 86 + $0.50". Then in the final experimental section, you would be be paid according to Contract 86. You would also receive an **additional** $0.50, along with your show up fee and any other earnings you receive for completing tasks. If you had instead selected "Contract 74", you would be paid according to Contract 74, and would not receive an additional payment.

- Each line of the question considers the same two contracts, only the additional payments change from line to line.
- Notice that the option on the right becomes more attractive as you move further down the question.
- Again, each line is a separate question, so you must make a selection on **EACH** line of the question.

Again, subjects in the random contracts version saw additional instructions:

Random Contracts:

In some of these questions, one or both of the possible options you are choosing between involve being randomly assigned to one of two contracts.

For example, you might see the question below:

Question 8: Please choose between the contracts below:

Contract 86

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.20
60+	0.20

Target: 20 tasks.
Reward: $0.20.
Consolation Prize Target: None.
Consolation Prize Reward: None.

or

Contract 82

50% Probability

Contract 74

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.40
60+	0.40

Target: 40 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

50% Probability

Contract 75

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.20
60+	0.40

Target: 60 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

○ Contract 86 + $0.50	○ Contract 82
○ Contract 86 + $0.15	○ Contract 82
○ Contract 86 + $0.10	○ Contract 82
○ Contract 86 + $0.05	○ Contract 82
○ Contract 86 + $0.01	○ Contract 82
○ Contract 86	○ Contract 82
○ Contract 86	○ Contract 82 + $0.01
○ Contract 86	○ Contract 82 + $0.05
○ Contract 86	○ Contract 82 + $0.10
○ Contract 86	○ Contract 82 + $0.15
○ Contract 86	○ Contract 82 + $0.50

In this question, one of the options, Contract 86, is a normal contract. The other option in this example, Contract 82, will randomly assign you to one of two contracts with equal probability.

If you choose this option on a question that is chosen to determine your contract in the next section, you will have a 50% chance of using the contract on the left, Contract 74, and a 50% chance of using the contract on the right, Contract 75.

Like when you were previously assigned a random contract, in order to determine which of the two contracts you would use, the computer will randomly generate a number between 1 and 100. If that number is between 1 and 50, you will use the contract on the left. If that number is between 51 and 100, you will use the contract on the right.

If you have chosen a random contract on the line that is selected to determine your contract, the random number will be generated **BEFORE** you begin completing tasks. This means that you will know which contract determines your payment before you begin completing any tasks.

Any additional payments that accompany an option are the same, regardless of which contract gets assigned. For example, suppose you had selected "Contract 82 + $0.05" on the line that gets chosen to determine your contract in the final section. You will receive the additional $0.05 regardless of whether you get assigned Contract 74 or Contract 75.

This means the order of events is:

- You make choices between the options of each line of the question.
- One of those questions is chosen at random, and one line of that question is chosen at random.
- Your choice on that line determines your contract in the final section. Any additional payments that accompany that choice are added to your earnings.
- If that choice involves a randomization between two contracts, the computer generates a random number to determine which of those contracts you will use, and tells you which contract was chosen.
- With that contract, you may complete as many or as few tasks as you wish, and receive a payment for those tasks according to the chosen contract.

At this point, subjects had an additional comprehension check to ensure that they understood the implementation of the random contracts.

Check your understanding:

Before you continue with the rest of the experiment, please answer the following questions to make sure that you understand the types of contracts you will be choosing between.

Below you will see an example of a contract that involves randomization. Please look at the contract and answer the following questions.

Contract 82

50% Probability	50% Probability

Contract 74

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.40
60+	0.40

Target: 40 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Contract 75

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.20
60+	0.40

Target: 60 tasks.
Reward: $0.40.
Consolation Prize Target: 20 tasks.
Consolation Prize Reward: $0.20.

Question 1: If this contract is the one selected for you to use in the final experimental section, will you:

- Use Contract 74 for sure.
- Use Contract 75 for sure.
- Use Contract 74 with probability 50% and Contract 75 with probability 50%.

Question 2: Which of the two contracts (74 and 75) are you more likely to be assigned?

- You are more likely to use Contract 74.
- You are equally likely to use Contract 74 or Contract 75
- You are more likely to use Contract 75.

Question 3: If this were used for the task section, when would you find out whether you had contract 74 or 75?

- You would find out whether you were going to use contract 74 or 75 **before** you begin working on tasks.
- You would find out whether you were going to use contract 74 or 75 **after** you begin working on tasks.

All subjects were given a practice version of the choice between menus to familiarize them with the interface:

Before you begin responding to the questions that will determine your contract for the final section, here is a practice question to familiarize you with the process of completing these questions and to demonstrate how your contract in the final section will be selected.

This question is like the ones that you will encounter afterward, except that it will **not** be selected to determine your contract in the final section.

As with the real questions, please select an option on each line of the question.

Practice Question: Please choose between the contracts below:

Contract 62		Contract 86	
Tasks completed	**Payment**	**Tasks completed**	**Payment**
0-19	0.00	0-19	0.00
20-39	0.00	20-39	0.20
40-59	0.40	40-59	0.20
60+	0.40	60+	0.20
Target: 40 tasks. Reward: $0.40. Consolation Prize Target: None. Consolation Prize Reward: None.		Target: 20 tasks. Reward: $0.20. Consolation Prize Target: None. Consolation Prize Reward: None.	

○ Contract 62 + $0.50	○ Contract 86
○ Contract 62 + $0.15	○ Contract 86
○ Contract 62 + $0.10	○ Contract 86
○ Contract 62 + $0.05	○ Contract 86
○ Contract 62 + $0.01	○ Contract 86
○ Contract 62	○ Contract 86
○ Contract 62	○ Contract 86 + $0.01
○ Contract 62	○ Contract 86 + $0.05
○ Contract 62	○ Contract 86 + $0.10
○ Contract 62	○ Contract 86 + $0.15
○ Contract 62	○ Contract 86 + $0.50

Subjects made a practice choice over these menus, and were then shown an example of how the computer would randomly draw one of those lines to decide which contract would be implemented in the fourth task section.

As described in the instructions, to determine which contract you will use in the final section, first one of the questions will be selected at random and then one of the lines from that question will be chosen at random.

As an example, suppose that the practice question was selected, and the row that was randomly selected was row 7. On that line, you chose Contract 62.

Contract 62

Tasks completed	Payment
0-19	0.00
20-39	0.00
40-59	0.40
60+	0.40

Target: 40 tasks.
Reward: $0.40.
Consolation Prize Target: None.
Consolation Prize Reward: None.

Contract 86

Tasks completed	Payment
0-19	0.00
20-39	0.20
40-59	0.20
60+	0.20

Target: 20 tasks.
Reward: $0.20.
Consolation Prize Target: None.
Consolation Prize Reward: None.

Contract 62 + $0.50	Contract 86
Contract 62 + $0.15	Contract 86
Contract 62 + $0.10	Contract 86
Contract 62 + $0.05	Contract 86
Contract 62 + $0.01	Contract 86
Contract 62	Contract 86
Contract 62	Contract 86 + $0.01
Contract 62	Contract 86 + $0.05
Contract 62	Contract 86 + $0.10
Contract 62	Contract 86 + $0.15
Contract 62	Contract 86 + $0.50

This means that if this were the real question, for the final experimental section you would be paid according to Contract 62.

Your earnings from the experiment would be the sum of your earnings in the task sections and your participation fee.

This concludes the practice question. The following questions will be used to determine the contract that you will use in the final section.

After this, subjects responded to the questions that asked them to choose between contracts, as shown in the main text. After answering all of those questions, they saw a screen informing them of the contract that had been selected for implementation, as in the example:

Selected Contract:

Question 2 was the question randomly selected for contract implementation, and from that question the selected row was 11.

In that round, you chose the following option: Contract #96

Contract 96

Tasks completed	Payment
0-19	0.00
20-29	0.20
30-49	0.20
50-69	0.20
70+	0.20

Target: 20 tasks.
Reward: $0.20.
Consolation Prize Target: None.
Consolation Prize Reward: None.

In the event that the selected contract was a random contract, the subjects saw the outcome of the random draw, and the implemented contract, similar to the style of the earlier task sections. After being told of the contract to be used in the fourth section, subjects completed the fourth and final task section.

Once finished with the final section, subjects were shown a screen reminding them of their earnings from each section, and their total payment, along with the randomly generated code used to claim their earnings through the MT interface. For example:

Individual Decision-Making Experiment

Finish

You have finished completing task sections.

For participating in the experiment, you earned $0.25.

For the sections you completed, you earned an additional:

In section 1 you earned $0.00.

In section 2 you earned $0.00.

In section 3 you earned $0.20.

In section 4 you earned $0.00.

Additionally you will receive the $0.15 bonus that accompanied your contract selection.

This means your total payment for this experiment will be $0.6.

Thank you for your participation.

Thank you for participating via Amazon's Mechanical Turk. In order to be paid your full amount for this experiment, you will need to enter the code found below.

Your bonus code is: **g3aVtEZLLQQiYpc**.

When you return to Mechanical Turk, please enter this code EXACTLY as it appears, and you will be paid the balance of your experimental earnings as your bonus for completing the task.