Impact on Financial Decisions through Resources and Education

Dr. Pallavi Mane[1], Worakamol Wisetsri[2], Shashi Kant Gupta[3]

[1]Assistant Professor, Indore Management Institute and Research Centre.

[2]Associate Professor, Department of Manufacturing and Service Industry Management, Faculty of Business and Industrial Management, King Mongkut's University of Technology, North Bangkok, Bangkok Thailand.

[3]Integral University, Lucknow, UP. India.

Abstract:

In high stakes financial decisions, people leave a substantial amount of money on the table, even when financial education is available. The ubiquity of financial choices makes understanding the effects of incentives and education on mistakes crucial. This research experimentally examines the impact of changes in incentives and educational availability on incentivized but hypothetical healthcare choices using Amazon Mechanical Turk. We find that increasing incentives are ineffective in increasing decision-making effort, even when these changes are made clear and salient to the subjects. Yet, surprisingly, despite this lack of effort response, subjects' choices improve when incentives are high. This result highlights an under-appreciated channel of incentives: when stakes become larger, often, the problems become simpler too. We next investigate the effect of available education. Overall, education leads to an increase in decision-making effort and an improvement in choice quality. However, this average effect masks significant heterogeneity across incentive treatments. Subjects are willing to put in the educational effort when either the problems are hard or mistakes are highly costly, but the return of the educational effort is zero for hard problems and positive for easy ones. Thus, only when stakes are high and the problem is easy, does education have an effect. These findings suggest that people can be encouraged to get education for high-stakes decisions, and policy-makers have a role in simplifying problems to translate the extra effort into better choices.

1. Introduction

Research in household finance has shown that consumers make financially sub-optimal choices, thus, leaving a substantial amount of money on the table in various types of financial decisions (loans in Agarwal et al. (2009), Bertrand and Morse (2011), insurance in Bhargava et al. (2015), Abaluck and Gruber (2011), and investment in Beshears et al. (2011)). One obvious explanation is that people lack the skills to make these decisions well. Yet, the literature finds that these mistakes persist even when financial education is available. A recent meta-analysis on financial literacy and financial behaviours finds that education has

surprisingly little effects on choices (Fernandes et al. (2014)). Why do consumers choose financially sub-optimal products even though mistakes are highly costly and education is freely available?

As a starting point to answering the question, this project designs an experiment to identify the effects of incentives and education on financial choices. This experiment mimics the choice of health insurance, a setting in which consumers incur significant losses and often misunderstand how the products work (Bhargava et al. (2015), Loewenstein et al. (2013)). We recruit Amazon Mechanical Turk (MTurk) workers to choose insurance plans for hypothetical scenarios [1]. Each scenario consists of deterministic health-care needs so that the total health costs of all listed plans are deterministic. The objectively correct choice is then the lowest-cost plan, and subjects receive a higher payment for choosing a lower-cost plan. Within this task, we vary the stakes and access to education. The variations allow us to test whether incentives and education matter for choices, the mechanisms by which they do so, and the effects of their interaction.

To vary the stakes, we design low- and high-incentive treatments. Each subject is randomized into one incentive level, which corresponds to a high or low cost of an average mistake. We make incentives higher by changing the premiums such that the variance of the total health costs increases. Because this increases the difference between the best plan and a randomly chosen plan is higher, mistakes become more costly [2].

We first look at the impact of incentives on how much effort subjects put into choosing a plan. When people make bad choices even with high stakes, there are two possible explanations. Either people do not increase effort with high stakes, or the extra effort is in vain. Using time spent on the task as a proxy for effort, we find evidence for the former hypothesis: subjects do not increase effort when incentives are higher. There is no effect of high stakes on time, although the best plan in the high-incentive treatment is worth two times that in the low-incentive treatment.

One possible reason why subject do not spend more time could be that they do not know the stakes. In our experiment, as in real-life insurance decisions, without calculating the variance of the total health costs, subjects may fail to realize how much their mistakes matter. If they do not know that they are in a high-stakes environment, they may not put in the effort. To test this hypothesis, within each incentive level, we implement another treatment, disclosure, in which subjects are told the stakes before they choose insurance plans. We find that disclosure does not change the results: knowing the underlying stakes does not impact the time spent. We, thus, conclude that subjects do not spend more time deciding because they perceive the returns of effort to be small, or at least smaller than the increase in incentives.

2. Related work

This paper is related to a number of strands of literature. We discuss them from the specific literature on health insurance to the broadest literature on the effects of incentives in experiments.

The relatively recent attention of US policy-makers on health insurance has been matched with a number of studies using US data, which generally show that consumers are not

choosing the financially optimal plan. Abaluck and Gruber (2011, 2016); Bhargava et al. (2015) use individual choices (under Medicare Part D for the first two papers, employer sponsored health insurance for the third paper) to show that consumers can save a significant amount of money by making a different choice. Moreover, Abaluck and Gruber (2016), which follow up on Abaluck and Gruber (2011), reveals that consumers do not make better choices over time. Besides these observational studies on poor choices, Loewenstein et al. (2013) presents evidence that consumers have little understanding of health insurance.

There are a few experimental papers trying to understand health insurance choices. Kling et al. (2012) uses a field experiment to measure the frictions in comparing insurance plans. Other papers have used hypothetical health choices (Johnson et al. (2013); Bhargava et al. (2015)). While our design is closest to that of Johnson et al. (2013), none of the existing papers manipulate incentives by changing the features of the plans or study the effects of incentives on effort and the interaction between incentives and education. By manipulating incentives via the premiums, our paper shows that high incentives do not affect effort, but they still matter by reducing difficulty. We also find an interaction between incentives and education on improving choices [3].

Bhargava et al. (2015) is the only paper we know that look directly at the effect of education on health insurance choices. Although their education treatment improves choices, this treatment is confounded by a comprehension test. Specifically, subjects who receive education are given a comprehension test before their choices while subjects who do not receive education are given the test after their choices. Besides, their experiments are not incentivized. In contrast, our experiment does not ask subjects for their understanding of the concepts before they choose a plan. We also use an incentivized setting which allows to study the interaction between education and incentives.

Moving beyond health insurance, our paper is nested within the disclosure and financial education literature. Existing papers disclose incentives by translating financial concepts (for instance, interest rate in the context of borrowing and saving) into dollar amounts (Bertrand and Morse (2011); Goda et al. (2014)). We disclose incentives by showing the dollar difference between the best option and a randomly chosen option. On financial education, there have been enough studies to prompt a meta-analysis by Fernandes et al. [4]. However, they have said little about the factors contributing the education effectiveness besides the field experiments by Drexler et al. (2014) and Carpena et al. (2017). As field experiments are limited by their ability to vary the choices, our experiment using hypothetical plans shed lights on how differences in the environment, such as premium change, can complement education [5].

On the broadest literature on the effects of incentives, there have been many experiments from the laboratory to the field on the effects of incentives (Camerer and Hogarth (1999), Gneezy et al. (2011)). The results are mixed and dependent on the types of tasks subjects complete [6]. Our experiment contributes to this literature using a hypothetical choice which mimics a real-life choice and shows that incentives may matter by making choices easier.

3. Framework

This section presents a framework capturing the key elements in the decision environment: incentive, disclosure, and education availability, which we map to the health insurance task. Then, using the framework, we show that measuring effort identifies the channels through which incentives and education have an effect.

Setting

Consider a decision-maker (DM) i who chooses from a list of insurance plans. This choice can be of high-stakes or low-stakes: the difference between the best plan and a random choice can be large or small. The stakes are denoted as an unknown $s \in \{H,L\}$ with a known prior $\mathbf{P}_{(s = H) = \mu}$.[8] These stakes can be disclosed or undisclosed. We denote disclosure as $d \in \{U,D\}$. Educational materials may or may not be available, denoted as $l \in \{0,1\}$.

The above three elements, s, d, and l, feature in the DM's timeline to choose a plan as follows:

1. The DM forms her belief of the stakes, $\hat{s}_{i(s,d) = \mathbf{P}i(s = H)}$

2. The DM decides how much effort, e_i, and how much educational effort, e_{il} to spend

3. The DM receives the result of her choice, $f_{i(ei;s)}$

We explain each of the above steps in turn. Before attempting the choice, the DM forms her belief of s, $\hat{s}_{i(s,d)}$. If she is in the disclosed treatment, $d_{= D}$, we display s transparently. As a result, the DM's belief is degenerate and correct: $\hat{s}_{i(H,D) = 1}$ and $\hat{s}_{i(L,D) = 0}$. If the DM is in the undisclosed treatment, $d_{= U}$, we do not give her any other information about s except the prior μ. She may examine and compare the plans to move her belief (correctly or incorrectly) towards either H or L. She may decide not to do so and maintain the belief at μ. In any of the cases, $\hat{s}_{i(\cdot)}$ is the DM's belief before she makes any decisions.

The DM then decides how much effort $e_{i \geq 0}$ to choose a plan. If education is available, $l_{= 1}$, then e_i may contain $e_{il \geq 0}$, the effort put into studying the educational materials [7].

Formally, we decompose e_i as $e_{i = eil + ein}$, where e_{il} is the educational efforts and e_{in} includes all other types of effort. Note that when education is unavailable, $l_{= 0}$, then $e_{il} = 0$.

The DM's effort translates to the number of correct answers, $f_{i(ei;s) = fi(eil,ein;s)}$ where $f_{i(eil,ein;\cdot)}$ is concave in each of the component of effort. We allow s to affect the number of correct answers (conditional on the same level of effort) because, under high stakes, the plan costs are further apart, so it may be simpler to tell a good plan from a bad one [8]. For example, if the DM estimates the costs of the plans, under high stakes, she needs to make a large

estimation error to confuse the relative quality of the plans. Meanwhile, under low stakes, she need only make a small error to choose the wrong plan. s can, thus, affect fi directly.

The DM's optimization problem is: maxeil,ein $\mathbf{E}s\hat{}i(\cdot)(s)fi(eil,ein;s) − (eil + ein)$ subject to e_{il} = 0 if l = 0. The DM perceives the return to her effort to be her task performance $f_{i(\cdot)}$ multiplied by the expected reward of doing well, which is the expected stakes under her belief $s\hat{}_{i(\cdot)}$. Since $f_{i(eil,ein,\cdot)}$ is concave, we can assume that the cost of effort is linear, without loss of generality [9]. Let e^*_i the DM's choice, which $\hat{s}_i(\cdot)(s)\dfrac{\partial f_i(e_i(\cdot); s)}{\partial e_i^*} = 1$ $(s,d,l) = (e^*_{il}(s,d,l),e^*_{in}(s,d,l))$ be satisfies$*$ Because we assume that $f_{(\cdot)}$ is concave, then an increase in $s\hat{}_{i(\cdot)}$, i.e., a greater belief that the choice is high-stakes, leads to an increase in $e^*_i(\cdot)$. Note that the choices of $e^*_{il}(\cdot)$ and $e^*_{in}(\cdot)$ satisfy the same condition. With this framework, we can study the effects of incentives and education.

4. Experimental setup

This section details our experimental design. We first outline the experimental setting: our subjects, their tasks, and their compensation [10]- [12]. We then describe the treatments in the main experiment and the follow-up experiment. Finally, we describe the data we collect for performance and effort.

4.1 Decision Environment: Subjects, Tasks and Payment

We recruited participants from Amazon Mechanical Turk (MTurk), which is a platform used by many social science experiments seeking a more representative population than university students. We restricted the subject pool to US workers because we would like the subjects to be familiar with the US health plan structure, which we use to design our plans. We posted the experiment as a Human Intelligence Task (HIT) [13]. Those who accepted the HIT followed a link to the experiment designed in Qualtrics, an online survey platform. For compensation, we paid them a participation fee of $2 and a bonus based on their choices in the experiment. The bonus is designed to incentivize subjects to spend effort, as further illustrated below.

As MTurk is an online platform, there is a worry that bots, instead of human workers, participated in the experiment. To minimize this concern, we restricted the subject pool to those who have completed more than 1,000 tasks and with approval ratings of more than 95%. Besides, workers needed to pass a captcha before entering in our experimental page [14]. Subjects completed two tasks: a calculation task, and then, a health insurance task.

Since we expect subjects to base their insurance choices on arithmetic estimations, we use the calculation task to understand the subjects' baseline motivation and skills. We also classify subjects based on subjects' performance in the calculation task to measure heterogeneous treatment effects [15]. At the end of the experiment, we choose one question from each task randomly and convert subjects' choices to the bonus.

We describe each task in turn. In the calculation task, there are four questions, an example of which is in figure 1. Each question contains four options, each option a sum. Subjects choose

one sum, which earns points amounting to 5,000 minus the chosen sum. For example, for the question in figure 1, the first sum is 4,880. If a subject picks this sum, her points are 5,000 – 4,880 = 120. The bonus payment is then 1 cent for each point. So, picking the first sum earns $1.20 if the question in figure 1 is chosen for payment. In this way, a subject earns the most points, and hence, the most money if she picks the smallest sum.

Figure 1: Calculation Task

After the calculation task, subjects complete the main task, a health insurance task. This task has five questions, an example of which is in figure 2. Figure 2 a zoom into the structure of the question. First, there is a hypothetical deterministic health-care scenario. Second, there are four plans whose structure mimics a US health insurance plan with a deductible, a co-payment/co- insurance, and a maximum out-of-pocket cost. Subjects choose a plan for the scenario. Because the health care scenario is deterministic, the costs of all plans are deterministic. The lowest-cost plan is the objectively correct answer. Subjects' points in this task equal 10,000 minus the total cost of the chosen plan. So, they are incentivized to choose the lowest-cost plan, which matches real-life decisions. This payment scheme explains the scheme for the calculation task: we would like to maintain consistency in how we pay subjects to minimize confusion [16].

Each question in the insurance task has accompanying materials to help subjects choose a plan. The materials always include glossary definitions which are the standard definitions available with any real-life plan. Figure 2b shows a complete screenshot of a question followed by the materials. At the end of the experiment, we ask subjects debriefing questions. For example, we ask subjects what they think is the stakes underlying the questions. Note that this debriefing happens before subjects know their final pay-out, so their answers are not affected by potential feedback.
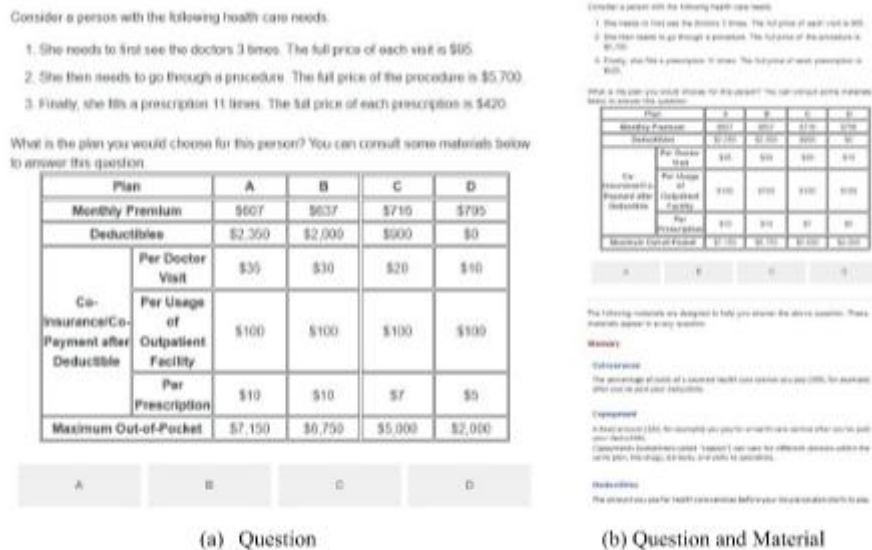
(a) Question (b) Question and Material

Figure 2: Health Insurance Task

4.2 Treatments in Main Experiment

The experimental treatments apply to only the health insurance task. To minimize confusion and spill overs across treatments, we use a between-subject design. Subjects are randomized into 2×2×2 (high versus low incentive × no disclosure versus disclosure × no education versus education) treatment cells summarized in table 1.

Table 1: Between-subject Treatment

		No Education	Education
LOW	Undisclosed	LU0	LU1
	Disclosed	LD0	LD1
HIGH	Undisclosed	HU0	HU1
	Disclosed	HD0	HD1

The health insurance task has three components corresponding to each treatment.

1. Incentive display: corresponds to disclosure treatment
2. Question: corresponds to incentive treatment
3. Accompanied materials: correspond to education treatment

As the question is the main component of the task, we explain this component first and then show how the other two components support answering the questions [17]. We vary the incentives in the questions by changing the premiums of the plans while keeping the scenario and all other features of the plans the same. Figure 3, highlights the monthly premium row, the only difference across incentive levels. There are two reasons to focus on altering premiums instead of other features to increase stakes [18]. First, changing the premium maintains the structure of the questions: the "out-of-pocket costs" of the plans are the same,

both in the amount and the calculation method. In other words, the more complicated part of finding the total health costs, which requires subjects to compare the health care needs with deductibles and co-payments or co-insurance, is the same across the incentives. The treatments differ in the simpler part: which number needs to be multiplied by 12 to find the yearly price of health insurance [19]. The second reason to use premiums is that when we survey the plans in the market, plans across companies are often the same in their features except for the premiums. As a result, we keep the same plan structure in the market and vary only the premiums.

Consider a person with the following health care needs.

1. She needs to first see the doctors 3 times. The full price of each visit is $95.
2. She then needs to go through a procedure. The full price of the procedure is $5,700.
3. Finally, she fills a prescription 11 times. The full price of each prescription is $420.

What is the plan you would choose for this person? You can consult some materials below to answer this question.

Plan		A	B	C	D
Monthly Premium		$607	$637	$716	$795
Deductibles		$2,350	$2,000	$900	$0
Co-Insurance/Co-Payment after Deductible	Per Doctor Visit	$35	$30	$20	$10
	Per Usage of Outpatient Facility	$100	$100	$100	$100
	Per Prescription	$10	$10	$7	$5
Maximum Out-of-Pocket		$7,150	$6,750	$5,000	$2,000

Consider a person with the following health care needs.

1. She needs to first see the doctors 3 times. The full price of each visit is $95.
2. She then needs to go through a procedure. The full price of the procedure is $5,700.
3. Finally, she fills a prescription 11 times. The full price of each prescription is $420.

What is the plan you would choose for this person? You can consult some materials below to answer this question.

Plan		A	B	C	D
Monthly Premium		$615	$645	$686	$811
Deductibles		$2,350	$2,000	$900	$0
Co-Insurance/Co-Payment after Deductible	Per Doctor Visit	$35	$30	$20	$10
	Per Usage of Outpatient Facility	$100	$100	$100	$100
	Per Prescription	$10	$10	$7	$5
Maximum Out-of-Pocket		$7,150	$6,750	$5,000	$2,000

(a) Low incentive   (b) High Incentive

Figure 3: Incentive Treatment

Note that our design of the incentive treatment differs from the standard method of manipulating incentives. In most decision experiments, incentives change because the exchange rate between experimental points and bonus changes (Johnson et al. (2013), Dewan and Neligh (2017)). For instance, 1 point can be converted to either 1 cent or 2 cents, and the 1-cent treatment is of low incentive. We use this exchange-rate method in our follow-up experiment. However, in the main experiment, we vary the incentives by changing the points of the plans and keeping the exchange rate constant because this is how plans are presented in the real world. In real-life choices, "exchange rate" is always the same as there is only one currency, but how much plans cost in that currency can change. When that happens, subjects can figure out the incentives, but they may not do so. As a result, the effects of incentives may be diminished because subjects do not know the stakes. By mimicking incentives in real-world decisions, we can then ask if undisclosed stakes affect choices less than disclosed stakes.

To vary disclosure, before the questions, we randomize subjects to see different screens informing them of the incentive levels. Under no disclosure, subjects see the prior

distribution (50% chance they are in either treatment). Under disclosure, subjects see the specific incentive to which they have been assigned. Figure 4 shows the difference across disclosure treatments conditional on the incentive level being low.

If you **choose a plan randomly**, you will receive **on average $2.25**.

In this section, there is a **50% chance** that you are in group X and a **50% chance** that you are in group Y.

If you are in **group X**, the **lowest-cost plan** gives you **on average $3.5**.

If you are in **group Y**, the **lowest-cost plan** gives you **on average $7**.

If you **choose a plan randomly**, you will receive **on average $2.25**.

If you always choose the **lowest-cost plan**, you will receive **on average $3.5**.

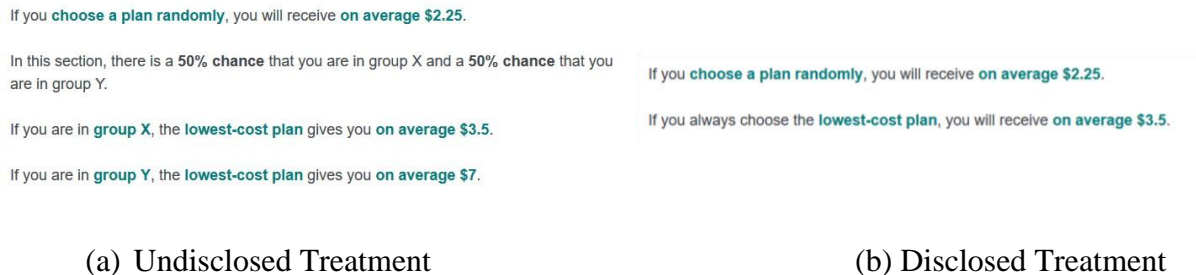(a) Undisclosed Treatment   (b) Disclosed Treatment

Figure 4: Disclosure Treatment, Conditional on Low Incentive

To vary available education, we randomize subjects to receive additional worked-out examples in the materials at the end of each questions. Although all subjects have access to glossary definitions, we note that this information does not show subjects the process of figuring out the right choice. In contrast, the examples show all the necessary steps to solve problems similar to those subjects have to answer. In other words, we devise deterministic health scenarios similar to those in the questions, and guide subjects on the cost calculation for two sample health plans. These extra materials are accessible via a series of buttons, so subjects can choose to use the examples or not.

4.3 Measurements

We measure choice by the number of questions subjects choose the best plan. To proxy for overall effort, we use the amount of time subjects spent choosing insurance plans because careful decisions take time. To proxy for educational effort, we use the time lapse between button clicks in the educational materials. We ignore the first click to minimize capturing impulsive clicking. If subjects click on the second button, we consider the subjects to have used the materials. We use the time taken between the second click and the last click within the education section to measure education time. Since subjects can continue reading or processing the materials after the last click, we note that our measure is the lower bound of the actual time spent on education. Besides, we recognize that time misses effort intensity. We provide suggestive evidence that this is not a significant concern in the results, and address this shortcoming more explicitly in the follow-up experiment.

In summary, we collect the following data for performance and effort:

- $f_{i(\cdot)}$: the number of questions where subjects choose the best plan

- e_i^*: the total time subjects spend on the insurance task

- e_{il}^*: click data within the education section

We use the above data and apply the analysis from the framework in next section to understand the effects of incentives and education in subjects' behaviours.

5. Results

All of the results in this section are from the main experiment described, although we interject with the results from the follow-up experiment where appropriate. Before discussing the treatment effects, we give an overview of the subjects by providing descriptive statistics of their demographics and their performance in the calculation task, which classifies them into types. We proceed to discuss the effects of incentives, disclosure, and education availability. We wrap up with the heterogeneous effects using the types defined by the calculation tasks and the types of mistakes subjects make.

5.1 Description of Subjects

The main experiment collects 2,009 complete responses, which are distributed approximately evenly across treatments. Each treatment has between 249 to 253 responses. On average, subjects completed the experiment in 1,413 seconds (24 minutes), earning $0.7 from the calculation task and $2.6 from the insurance task (on top of the $2 participation fee).

There are no obvious concerns about selection bias. First, we randomly assign subjects into treatments. Second, we do not find attrition bias. Of the 123 incomplete responses, 90% abandon the experiment before the insurance task. The remaining 10% are present in all treatments. Third, we check for demographic balance across the treatments by checking for the treatment "effect" on the demographics, the result of which is in A.1. The only significant difference is that subjects randomized into the high-incentive treatment are less likely to have a health insurance plan, which is consistent with a 5% random chance of finding a significant difference.

There are three differences in demographics between the sample and the US population worth noting. First, the sample is more educated, with 57% having a college degree or more, compared to 31% in the US Census. Second, they are younger: there are relatively few workers beyond the age of 40. Third, more of the sample, 18%, do not have health insurance compared to the 9% in the population. The last two differences agree with our prior of "gig workers" on an online platform. That the sample is relatively young possibly explains their higher education level. Although the differences with the US population are not essential to the study per se, it is useful to keep in mind that this sample is not representative of the consumers, and our results are local to this population.

5.2 Calculation Task

We give an overview of the subjects' performance on the task and then use their performance to classify them into two types. On average, subjects spend 4.1 minutes on the calculation questions, answering 1.94 questions correctly (out of 4). Although this performance is significantly better than randomization, at 1 correct answer, recall that the calculation task asks straightforward arithmetic questions, which subjects can complete however they wish, without a time limit. So, even when subjects can answer the questions perfectly, the cost of doing so is non-trivial.

Figure 4 presents the distribution of subjects' performance, which shows a fair amount of heterogeneity. The distribution is spread out over all the possible number of correct answers, from 0 to 4 possible correct answers. The vertical line, at 1.94, indicates the average number of correct answers. As a simple classification of subjects, we label those who answer more than 1 question correctly, corresponding to doing better than randomly, as the "high type", θ_h, and the rest as θ_l. In our data, this classification happens to be a reasonably even split with 53% of subjects belonging to θ_h and 47% belonging to θ_l. We use this classification to understand heterogeneous effects in the main task.

To provide a complete picture of the types, figure 5 presents the CDF of time θh and θl spend in different components of the experiment. Note that all analysis for time is done in logarithmic to correct for the heavy right tail in the time data. Figures 5a and 5b show that θh generally spend more time to answer the questions in both tasks.
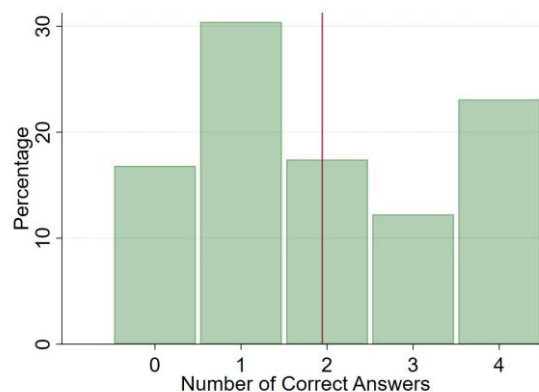

Figure 4: Distribution of Subjects' Performance in Calculation Task

Although that θl not spending time and not doing well may trigger the worry that they are bots spamming MTurk, we do not think this is the case. Figure 5c shows that there is no difference between the types in the time spent outside of the tasks, i.e., reading instructions. So, θl spend time to read instructions to understand the experiment, but when the questions appear, they do not spend time and answer them poorly.

5.3 Insurance Task
We give an overview of the subjects' performance before discussing the effects of incentives and education. On average, subjects spend 335 seconds (5.6 minutes) on the insurance task, answering 1.6 questions correctly (out of five). There are variations in performance: subjects in LU0 treatment have the worst performance with 1.43 correct answers, and those in HD1 perform the best with 1.89. The performance in all treatments are significantly better than 1.25, the average accuracy under randomization. 33% of subjects use the materials when they are available although they seem to spend a limited amount of time, only about half a minute, looking at the materials.
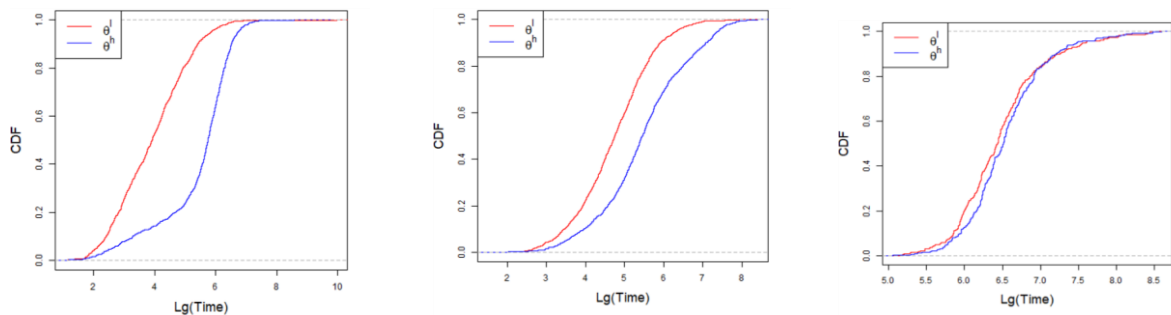
(a) CDF of l_g in Calculation Task   (b) CDF of l_g in Insurance Task   (c) CDF of l_g in Non-Task

Figure 5: Overall Time Spent in Components of the Experiment by Subjects' type

6. Conclusion

Motivated by the poor choices in consumer finance even though stakes are high and education is freely available as documented by literature, this project implements an experiment on MTurk to study the effects of incentives and education on behaviours. We use the health insurance setting and vary stakes underlying the decisions, information about the stakes via disclosure, and access to education. To pin down the mechanisms of incentives and education, we measure both choices of insurance plans and time spent in the task.

There are three main findings of the experiment. First, high incentives have a surprising alternative channel in making the problems simpler. Second, subjects perceive education to be beneficial in either hard problems or high-stakes environment. Third, the actual return of education is positive for easy problems but zero for hard problems. The combination of the last two findings suggests that in financial choices, people could be encouraged to use education when stakes are high, but policy-makers should aim to reduce the difficulty of the problems so that the educational effort translates to better choices.

Overall, even with our full intervention of high incentives and available education, the average performance is still poor. Subjects answer fewer than half the number of questions correctly. As a result, there is much space for future research to understand choices and explore potential solutions.

References

[1] Abaluck, J. and Gruber, J. (2011). Heterogeneity in choice inconsistencies among the elderly: Evidence from prescription drug plan choice. American Economic Review, 101(3):377–81.
[2] Abaluck, J. and Gruber, J. (2016). Evolving choice inconsistencies in choice of prescription drug insurance. American Economic Review, 106(8):2145–84.
[3] Afzal, U., d'Adda, G., Fafchamps, M., Quinn, S., and Said, F. (2018). When nudge comes to shove: Demand for commitment in microfinance contracts. mimeo.

[4] Agarwal, S., Skiba, P. M., and Tobacman, J. (2009). Payday loans and credit cards: New liquidity and credit scoring puzzles? American Economic Review, 99(2):412–17.

[5] Agranov, M. and Ortoleva, P. (2017). Stochastic choice and preferences for randomization.

[6] Journal of Political Economy, 125(1):40–68.

[7] Agranov, M. and Ortoleva, P. (2020). Ranges of preferences and randomization. mimeo.

[8] Ahn, D. and Sarver, T. (2013). Preference for Flexibility and Random Choice. Econometrica, 81(1):341–361.

[9] Amador, M., Werning, I., and Angeletos, G.-M. (2006). Commitment vs. Flexibility. Econometrica, 74(2):365–396.

[10] Ambuehl, S., Bernheim, B. D., and Lusardi, A. (2014). Effect of financial education on the quality of decision making. National Bureau of Economic Research.

[11] Andersen, S., Campbell, J. Y., Nielsen, K. M., and Ramadorai, T. (2015). Inattention and inertia in household finance: Evidence from the danish mortgage market. National Bureau of Economic Research.

[12] Apesteguia, J. and Ballester, M. A. (2018). Monotone stochastic choice models: The case of risk and time preferences. Journal of Political Economy, 126(1):74–106.

[13] Ashraf, N., Karlan, D., and Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the Philippines. The Quarterly Journal of Economics, 121(2):635–672.

[14] Bernheim, B. D., DellaVigna, S., and Laibson, D. (2019). Handbook of Behavioral Economics Foundations and Applications. Elsevier.

[15] Bertrand, M. and Morse, A. (2011). Information disclosure, cognitive biases, and payday borrowing. The Journal of Finance, 66(6):1865–1893.

[16] Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2009). The importance of default options for retirement saving outcomes: Evidence from the united states. In Social security policy in a changing environment, pages 167–195. University of Chicago Press.

[17] Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2011). How Does Simplified Disclosure Affect Individuals' Mutual Fund Choices? pages 75–96. University of Chicago Press.

[18] Bhargava, S., Loewenstein, G., and Sydnor, J. (2015). Do individuals make sensible health insurance decisions? evidence from a menu with dominated options. Technical report, National Bureau of Economic Research.

[19] Blackwell, D. (1953). Equivalent comparisons of experiments. The annals of mathematical statistics, pages 265–272.