7-2021

Information avoidance and medical screening: A field experiment in China

Yufeng LI
Capital Medical University

Juanjuan MENG
Peking University

Changcheng SONG
Singapore Management University, ccsong@smu.edu.sg

Kai ZHENG
University of California, Irvine

Management Science (forthcoming)

Information Avoidance and Medical Screening: A Field Experiment in China[1]

Yufeng Li[1], Juanjuan Meng[2], Changcheng Song[3], and Kai Zheng[4]

May 2020

Abstract: Will individuals, especially high-risk individuals, avoid a disease test because of information avoidance? We conduct a field experiment to investigate this issue. We vary the price of a diabetes test (price experiment) and offer both a diabetes test and a cancer test (disease experiment) after eliciting participants' subjective beliefs about their disease risk. We find evidence that, first, some people avoid the test even when there is neither a monetary nor transaction cost; and second, both low- and high-risk individuals select out of the test as the price increases. We explain our findings using three major models of anticipatory utility.

JEL Classification: D84, I12

Key Words: Anticipation Utility; Information Avoidance; Health Anxiety; Health Screening

[1] We would like to thank Hunt Allcott, Daniel Benjamin, Eric Finkelstein, Teck-Hua Ho, Botond Kőszegi, Agus Salim, Shyong Tai, and participants in seminars at the National University of Singapore, NBER Summer Institute Healthcare 2018, and Tsinghua BEAT 2018 for helpful comments. This work was carried out with the aid of a research grant from the NUS Initiative to Improve Health in Asia (NIHA) coordinated by the Global Asia Institute of the National University of Singapore and supported by the Glaxo Smith Kline-Economic Development Board (Singapore) Trust Fund. Juanjuan Meng acknowledges financial support from the National Natural Science Foundation of China (Grant No. 71103003 and Grant No. 71471004); Beijing Higher Education Young Elite Teacher Project (Grant No. YETP0040); (Solareast Corporation) Guanghua Research Fellow; and the Key Laboratory of Mathematical Economics and Quantitative Finance (Peking University). We thank Xinyao Qiu and Fuyang Zhao for their excellent work as research assistants. We also thank Jianghan Gu, Weisi Wang, Weiting Yuan, and Yufei Zhang for their efforts in collecting the survey data in China, and Xiangshuang Kong in particular for his help in facilitating the field work.

[1] Capital Medical University Beijing, Pinggu Hospital; email: doctorlyf@126.com
[2] Guanghua School of Management, Peking University; email: jumeng@gsm.pku.edu.cn
[3] Lee Kong Chian School of Business, Singapore Management University; email: ccsong@smu.edu.sg,
[4] Department of Informatics, University of California, Irvine; email: kai.zheng@uci.edu

1. Introduction

Information is valuable in standard economic analysis, because it improves decision-making. However, there are many situations in which people avoid useful information (see Golman et al. (2017) for a literature review). For instance, many empirical studies find that people tend to avoid important information regarding their health status, such as by declining to take a test or learn the results (Lyter et al. 1987; Lerman et al., 1999; Sullivan et al., 2004; Thornton 2008; Oster et al. 2013; Ganguly and Tasoff, 2016), and this tendency can lead to a lack of proper treatment for the disease. Most previous studies focus on the overall information avoidance effect; only recently have researchers begun to examine the heterogeneity effect. For instance, some studies examine whether high-risk individuals are more likely to avoid medical tests (Caplan, 1995; Nosarti et al., 2000; Meechan et al., 2002; Oster et al. 2013; Persoskie et al., 2014). Overall results are quite mixed, however, probably due to the absence of a clean identification strategy. Only a few studies examine how a tendency to avoid information varies with the test price (Thornton, 2008; Okeke et al., 2013) or type of disease (Ganguly and Tasoff, 2016).

This paper presents a randomized field experiment that examines the medical test take-up decision as a function of subjective disease risk, with exogenous variations on price and disease type. We find evidence for the information avoidance phenomenon whereby some individuals avoid taking the test for two diseases—diabetes and cancer—even when there is neither monetary nor transaction costs of taking the test. Our field experiment also shows that both low-risk and high-risk individuals are more likely to opt out of the test when the price is higher. These findings are puzzling from a neoclassical perspective, because a simple neoclassical model, with the assumption that treatment of the disease can only be implemented after being formally diagnosed, would predict a 100% take-up rate if there is no cost, and high-risk individuals should be more likely to take the test because of the larger benefit of obtaining information and, in turn, proper treatment.

Specifically, we collaborate with a local hospital to conduct a randomized field experiment with approximately 1,200 individuals in rural China. The field experiment has two designs: a price experiment and a disease experiment. In the price experiment, we vary the price of a diabetes test. Individuals were randomly assigned to one of three groups: the free group (T0), the 10 RMB group

(T10), or the 30 RMB group (T30).[1] In the disease experiment, individuals were randomly assigned to one of two groups: the diabetes group or the cancer group. We provided the disease test for free after blood had been drawn for another free blood test (so there was no additional cost for taking the test), but varied the disease type to be tested. In all treatments, we elicited individuals' self-reported beliefs about their corresponding disease risk before they made their testing decision so that we could investigate the heterogeneous effect across disease risk.

We define a simple neoclassical model as a benchmark case, which assumes that treatment of the disease can only be implemented after being formally diagnosed. This model would predict a 100% take-up rate if there is no cost of taking the test. We find that in the disease experiment, when there is neither a monetary cost nor a transaction cost (because the blood has already been drawn), about 14% and 11% of participants avoid the test for diabetes and cancer, respectively. In the price experiment with a free test, when there is only the cost of having blood drawn (i.e., the monetary cost of the test is zero), about 34% of participants avoid the test. We show that even with no monetary cost and little transaction cost, test avoidance still exists. This is the first empirical contribution of the paper.

A simple neoclassical model would predict that individuals with higher subjective beliefs about disease risk should be more likely to take the disease test, because the test outcome will allow them to undertake proper treatment—and since no treatment can be done without a diagnosis, the information is more valuable.[2] This view also suggests that as the test price increases, only high-risk individuals would remain in the testing group; hence the average test outcome in the testing group should be more severe as the test price increases. However, our cross-treatment results from the price experiment suggest that there is no significant difference in the mean value of blood glucose levels across treatments in the sample that took the diabetes test. More interestingly, the distribution of blood glucose levels becomes significantly less dispersed when the test price increases. The cross-treatment distribution suggests that as the test price increases, both high- and low-risk individuals select out of the test. Our cross-treatment results in the disease experiment also show that the tendency to avoid the test is more salient when the disease is more severe, and especially when individuals believe the disease is less controllable. These results are well identified, because prices and disease types are exogenously

[1] 1 USD=6.6 RMB in October 2017. A test price of 30 RMB is comparable to the market price.
[2] If we assume that individuals can also take other actions of the same treatment quality without undergoing medical screening, then the neoclassical model predicts that high-risk individuals will be less likely to take the test. See Section 4.1 for a detailed discussion.

varied on random samples of participants.

The within-treatment analysis also provides consistent evidence: The increase in test take-up with respect to subjective risk is decreasing in subjective risk in T30. Further regression analysis of the alternative channels suggests that these patterns are not driven by subjective risk being correlated with knowledge of the benefits of testing, compliance costs for undergoing treatment, or financial constraints from undergoing treatment if diagnosed as having diabetes.[3] When the cancer test is provided in the disease experiment, we find that those with either low or high subjective risk are less likely to take the cancer test. This tendency is significantly stronger when people believe that cancer is less controllable. We also find that the tendency for the take-up rate to increase with subjective risk is significantly weaker in the cancer treatment than the diabetes treatment.

To the best of our knowledge, this is the first experimental study from the field that provides both cross-treatment and within-treatment evidence that high-risk individuals may be more likely to avoid medical tests. This is the second empirical contribution of our paper.

We also find an interesting heterogeneous effect that deepens our understanding of when we can empirically observe the tendency for high-risk individuals to avoid the test. In T0 and T10 of the price experiment, when the test price is low, we find no concave relationship between the probability of taking the test and the subjective risk of having the disease. Similarly, when the diabetes test was provided for free in the disease experiment, although there is weak evidence that the high-risk group tends to avoid the test, this effect is not significant. In general, the effect of information avoidance on the high-risk group is more salient when the test price is high and when the disease is more severe. This heterogeneous effect is the third empirical contribution of our paper.

What models of information avoidance can best explain our findings? We illustrate the predictions from three major models of information avoidance regarding the relationship between the probability of having the disease and the net value of testing, which is often not explicitly stated in previous studies: the model of optimal expectations (Brunnermeier and Parker, 2005; Oster et al., 2013); the attention model (Karlsson, Loewenstein, and Seppi, 2009; Golman and Loewenstein, 2015; Ganguly and Tasoff, 2016); and the curvature model (Caplin and Leahy, 2001; Caplin and Eliaz, 2003; Kőszegi, 2003; Eliaz

[3] There are two other alternative explanations for the low take-up rate for medical tests. One, based on the neoclassical model, is the high price elasticity (Thornton, 2008); another behavioral explanation is procrastination generated by present bias. However, neither of the two explanations would predict that high-risk individuals would be more likely to avoid the test. Our experimental design also excludes procrastination, because all participants had already paid the upfront cost of showing up.

and Spiegler, 2006). All three groups of models assume anticipatory utility—i.e., beliefs about the future affect current utility. The models differ, however, in how anticipatory utility enters the utility function and whether beliefs are allowed to be irrational. We show that under the assumption that some treatment is possible without the test, all three groups of models can explain the two empirical findings: Some individuals avoid the test when the test has neither a monetary nor transaction cost, and both low- and high-risk individuals tend to select out of the test as the price increases. The antipatory utility ingredients in these models help in explaining information avoidance, but the assumption that treatment is possible without the test helps to generate the non-monotonic relationship between disease risk and take-up decision. These models are also able to explain the observed heterogeneity with price and disease type. However, even under the ancilliary assumption about treatment, the neoclassical model cannot explain information avoidance when the test is free.

This paper is related to both empirical and theoretical studies on information avoidance. Golman et al. (2017) provide an excellent review of this literature. As noted above, many empirical studies find that people tend to avoid important information regarding their health status (Lyter et al. 1987; Lerman et al., 1999; Sullivan et al., 2004; Thornton 2008; Oster et al. 2013; Ganguly and Tasoff, 2016). For instance, participants in Thornton's (2008) study generally avoided learning their HIV test outcomes, but even small incentives reduced the avoidance rate significantly. Most previous studies focus on the overall information avoidance effect; only a few investigate the heterogeneous effects across the probability of having the disease. The latter group produces mixed results; some find that people with higher risk of having cancer tend to delay a visit to the doctor (Caplan, 1995; Nosarti et al., 2000; Meechan et al., 2002; Persoskie et al., 2014). In contrast, using elicited subjective beliefs, Oster et al. (2013) find that individuals with higher subjective belief about disease risk were more likely to pursue being tested for Huntington's disease, and people were generally overly optimistic about the risk of having such disease. Okeke et al. (2013) conducted a randomized trial in Nigeria with varying prices for cervical cancer screening. Despite the lack of statistical significance, they found that high-risk subjects (for both subjective and objective risk) tended to accept a higher test price in general. To our best knowledge, our paper is the first field experiment to find that individuals with high subjective belief about their disease risk tend to avoid testing. We also identify some conditions from both our empirical and theoretical work under which this effect is more likely to appear, which helps to reconcile

the mixed results in the literature.

In terms of cross-disease comparison, Ganguly and Tasoff (2016) find that more people are willing to forgo a $10 payment to avoid learning the results of the herpes simplex virus 2 (HSV-2) test than an HSV virus 1 (HSV-1) test; HSV-2 is viewed as a more serious condition. Our comparison of diabetes and cancer yields a similar result. However, there are two differences between the two papers. First, we draw blood from all participants in the disease experiment before asking whether they want to know the results; Ganguly and Tasoff elicit the information preference using an incentive-compatible strategic method before drawing blood. Second, we conduct a between-subjects test and Ganguly and Tasoff a within-subjects test.

Three types of belief-based utility models can help to explain information avoidance in a medical testing context: the optimal expectations model (Brunnermeier and Parker, 2005; Oster et al., 2013); the attention model (Karlsson, Loewenstein, and Seppi, 2009; Golman and Loewenstein, 2015; Ganguly and Tasoff, 2016); and the curvature model (Caplin and Leahy, 2001; Caplin and Eliaz, 2003; Kőszegi, 2003; Eliaz and Spiegler, 2006). All three groups of models assume anticipatory utility—i.e., beliefs about the future affect current utility. Models differ, however in how anticipatory utility enters the utility function and whether beliefs are allowed to be irrational.

Both the model of optimal expectations and the attention model assume that individuals derive anticipatory utility from beliefs about future health status.[4] The optimal expectations model allows for self-manipulation of beliefs: Individuals avoid the test, since they can maintain biased beliefs to generate high anticipatory utility. In contrast, the attention model maintains the assumption of rational beliefs. In the attention model, a boost of anticipatory utility is felt immediately if a person receives information that is utility relevant to future experience. Individuals avoid the test, since they might receive a negative shock of anticipatory utility if they take it. The curvature model also maintains the assumption of rational beliefs: Individuals avoid the test because it increases the uncertainty of beliefs, which increases anxiety when the utility over beliefs is concave.

The paper proceeds as follows. Section 2 introduces the experimental design for the field study, and Section 3 presents empirical results. Section 4 presents the predictions for the three groups of

[4] Information avoidance in the setting of self-confidence can also be explained by the model of self-deception with endogenous memory (Bénabou and Tirole 2002). In this model, the agent weighs the benefits of preserving his effort motivation against the risk of becoming overconfident, and might choose to avoid bad news to conserve the self-confidence necessary to motivate acting.

information avoidance models, and Section 5 concludes.

2. Experimental Design

2.1. Background

As of 2016, 422 million people have diabetes worldwide, up from 108 million in 1980.[5] The prevalence of diabetes is 8.5% among adults—nearly double the rate of 4.7% in 1980 (WHO, 2016). Approximately 673 billion USD were spent on diabetes, which accounts for about 12% of global health expenditure (International Diabetes Federation, 2015). Many people remain undiagnosed, because often there are few symptoms during the early years of type 2 diabetes. About 46.5% of people with diabetes worldwide do not know they have the disease (International Diabetes Federation, 2015), and the number is higher in Asian countries. For example, 9.7% of the adult population in China has diabetes, and 60.7% of Chinese with diabetes do not know they have the disease (Yang et al., 2010). This lack of knowledge generates a huge cost: Diabetes mellitus caused 1.6 million deaths in 2015, making it the sixth leading cause of death (WHO, 2017).

Screening is potentially an important strategy to mitigate the effects of diabetes, since early detection and prompt treatment may reduce the burden of diabetes and its complications. Screening typically involves drawing venous blood to measure blood sugar and glycated hemoglobin. We offer the following types of blood tests for diabetes: random plasma glucose (RPG), fasting plasma glucose (FPG), and oral glucose tolerance (OGTT). The RPG consists of a blood check at any time of the day that does not require fasting, but is also not very accurate for diagnosing diabetes compared with the other two. The FPG requires fasting for at least 8 hours before the test. The 2-hour OGTT, which checks blood glucose levels before and 2 hours after drinking a solution of glucose and water, reveals how the individual processes glucose.[6]

We also included one test related to cancer in our study. The carcinoembryonic antigen (CEA) blood test is commonly used to follow patients with known cancers. It can also be used as

[5] Diabetes mellitus is a group of metabolic diseases in which high blood sugar levels are present over a prolonged period. The chronic hyperglycemia of diabetes is associated with long-term dysfunction, damage, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels.

[6] Belief-based models may predict particular preferences for the resolution of uncertainty. For instance, decision makers in Kőszegi and Rabin's (2009) study preferred quicker resolution of uncertainty—i.e., more accurate information. Since the RPG is less accurate than the other two, one might wonder whether the choice of test helps to distinguish various models. However, this is unlikely in our setting, because participants were not given information about accuracy, and the costs of different tests also differed.

a tumor marker, especially for cancers of the gastrointestinal tract. A rising CEA level is correlated with progression or recurrence of the disease. Note that the CEA by itself is not specific enough to substantiate the recurrence of cancer; further tests are required for confirmation. Details of these tests are provided in Appendix 1.

2.2. Experimental Design

We collaborate with a large local hospital in a rural county in Beijing, China, to study demand for these disease tests. The collaboration offers two advantages. First, doctors and nurses from the hospital can provide medical knowledge, medical tests, and related services. Second, the hospital can help us earn the trust of residents, which is necessary in order to conduct the study.

In 2014, ten villages in the county were randomly selected. We first collected administrative data—name, gender, birthdate, and address—from the local government for all individuals in the sample villages. We asked village leaders to instruct everyone who did not have diabetes to come to the village office on the day of the study, which allowed us to survey the full sample of eligible individuals. Upon arrival at the study site, we asked households to complete a survey in a separate room. We provided a free basic medical examination for all participants after the survey, which included height, weight, and blood pressure.

We designed two experiments to investigate what determines demand for diabetes screening: a price experiment and a disease experiment.[7] We conducted the price experiment in five villages and the disease experiment in the other five. Randomization is at the individual level to increase the power. Figure 1 presents the experimental design. In the price experiment, we varied the price of the diabetes test. When individuals arrived for the study, enumerators first administered the survey. Participants were randomly assigned to one of three groups *after* completing the survey: the free group (T0), the 10 RMB group (T10), or the 30 RMB group (T30). Individuals chose one of three sealed envelopes offered by enumerators, and the voucher inside the envelope stated the price they would have to pay to receive the diabetes test. The actual price to conduct the diabetes test in the hospital used for the study is 30 RMB. We then asked whether they would like to take a diabetes test. If so, nurses from the

[7] On the survey, we asked about health insurance status; about 96% had access to healthcare. This is because China established the National Cooperative Medical System (NCMS) in 2003, which provides rural residents with large subsidies from central and local governments.

local hospital drew their blood after the physical examination. We chose diabetes tests that use venous blood to measure blood sugar and glycated hemoglobin, which requires laboratory analysis and produces results several days later. If individuals had eaten breakfast before taking the blood test, we drew blood once and measured the random blood sugar level. If they had fasted before the blood test, we conducted the fasting blood sugar test or the oral glucose tolerance test, depending on the individual's choice.

[Figure 1]

In the disease experiment, we varied the disease being tested for after blood had been drawn. Village leaders informed all participants that there would be a free blood test to obtain basic blood counts and that they should fast before coming to the study. When they arrived, nurses first drew venous blood from all individuals and enumerators administered the survey. Participants were randomly assigned to one of two groups: the diabetes group or the cancer group. Randomization was conducted by the researcher using a computer, and individuals were not aware of their assignment. In the diabetes group, after drawing the blood and administering the survey, we asked whether participants would like to use the blood that had been drawn for an additional free diabetes test (fasting blood sugar). The procedure was the same for the cancer group, except we asked whether they would like to have an additional free test for cancer risk (CEA).[8] Participants in both groups were told that if they chose to have the additional test, nurses would send their test results via text message several days later.

We are interested in (1) the impact of different treatments on take-up of the screening test and (2) who selected to be screened under different treatments. The key information necessary to understand question (2) is diabetes risk, which can be determined by both objective and subjective measures. The subjective measure is self-reported beliefs about diabetes risk and cancer risk. We asked participants the following question: "What do you think is the probability that you have diabetes/cancer?" To indicate their answers, participants were given 10 small paper balls and asked to distribute them across two areas: (1) No diabetes/cancer and (2) have diabetes/cancer. If participants put 8 paper balls into (1) and 2 paper balls into (2), the perceived probability that they have diabetes/cancer is around 20%. Objective measures include test outcomes (which are only available for those who take the test).[9]

[8] The price of the CEA test in the same hospital is 40 RMB.

[9] In theory, one can predict diabetes risk from health measures such as BMI, blood pressure, and smoking habits, but such predictions

The survey also collects information on the participant's socioeconomic background, lifestyle, knowledge about diabetes, risk attitudes, time preference, and information avoidance. Appendix 2 presents all survey questions, and Table A1 in Appendix 4 explains how the variable was constructed for analytic purposes.

3. Experimental Results

3.1. Summary Statistics

We surveyed 664 individuals, with a response rate of about 93%, in the price experiment and 531 individuals, with a response rate of about 96%, in the disease experiment. The high response rate is due to the free medical examinations and high trust in village leaders and the local hospital. We begin by performing randomization checks across the price experiment and the disease experiment. Table 1 reports the mean and standard deviations of four groups of variables: screening decisions, demographic information, health conditions and behaviors, and preference measures. We use asterisks on the T30 variable to indicate whether the variables in T0, T10, and T30 are significantly different in the multivariate test. We use an asterisk to indicate whether variables in the cancer treatment are significantly different from those in the diabetes treatment.

[Table 1]

Panel A is the key decision variable: the take-up rate of tests in the treatment. Not surprisingly, as the price of the diabetes test rises, the take-up rate of the test declines significantly, from 0.66 (T0) to 0.37 (T10) and then 0.20 (T30). However, take-up rates in the disease experiment are 0.86 and 0.89 for diabetes and cancer tests, respectively—which are not significantly different from each other—and are much higher than in the price experiment. This is expected, because in the disease experiment both tests were free and individuals were asked whether they would like to take the test *after* their blood samples had been collected; as a result, the cost of taking the test is much lower.

Panel B to Panel D reports demographic information, health information, and subjective beliefs

---

are likely to be highly inaccurate due to the relatively small sample and the fact that diabetes is also affected by lifestyle, which is difficult to assess precisely. We tried to estimate the relationship between various health/demographic measures and the tested outcome using the disease experiment, performing out-of-sample prediction using the diabetes treatment, and finally by comparing the predicted level with the actual level to test the accuracy of the prediction. We also tried several specifications to predict objective risk by using (1) a logistical model to predict actual discrete test outcomes, (2) an OLS model to predict blood sugar levels, and (3) machine learning (Lasso and Ridge) to predict blood sugar levels. Under all specifications, estimation R squares are low and prediction R squares are even negative, suggesting that the predicted objective risk matches the actual outcome much less often than the sample average. It is not sensible to conduct any analyses with such poor prediction.

about disease risk. The average subjective assessment of disease risk is about a 10%-13% chance of getting diabetes and 10% of getting cancer. Figure A1 shows the distribution of subjective beliefs for both the diabetes treatment and the disease experiment. We have two observations. First, the majority of subjects report that their subjective risk of diabetes or cancer is zero. Second, there is large heterogeneity in their subjective beliefs about diabetes or cancer risk. One might wonder whether the self-reported subjective risks contain any real information. We show via regression that people who are better informed about diabetes, who are less able to follow treatment requirements, and who are more anxious in general have higher reported subjective risk of diabetes. Conditional on taking the diabetes test, the correlation between subjective beliefs and test outcome is 0.2188. Both types of evidence suggest that our self-reported beliefs contain relevant information. Overall, seven out of 72 tests from Panel B to Panel D are significant, which is expected under random assignment.

3.2. Price experiment

3.2.1. Cross-treatment results

In this section we analyze the cross-treatment pattern in the price experiment. Because price is exogenously varied, this part of the result is well identified. We begin by providing summary information for the diabetes tests. We asked individuals to fast before coming to our study. For those who were in a fasting state, the fasting plasma glucose (FPG) test was preformed. Ninety-two individuals took the FPG and were diagnosed as having diabetes if the outcome level exceeded 7 mmol/L.[10] For those who were not fasting, the random plasma glucose (RPG) test was performed. In total, 146 individuals took the RPG.

Figure 2 displays the take-up rate of the diabetes test across treatments. Not surprisingly, the take-up rate steadily declines as the price of the test increases. More than 60% of participants take the test when it is free, but this rate drops to about 40% when the price is 10 (T10), and to 20% when the price increases to 30 (T30). These changes are all statistically significant.

[Figure 2]

The simple neoclassical model that assumes treatment of the disease can only be implemented

[10] Of the 92 participants, 33 were willing to wait for 2 hours and take the OGTT, which requires two blood tests. The first is the same as the fasting plasma glucose test, and the second test is taken 2 hours after drinking a mixture of glucose and water. We use the first test results for these 33 participants for analysis, which yields the same diagnostic outcome as using the results from both tests.

after being formally diagnosed makes the following prediction: High-risk individuals are more likely to take the test. As a result, when the test price increases, high-risk individuals remain as test takers, while individuals with lower risk tend to select out of the test. Therefore, the average test outcome for those who take the test should report more diabetes risk as the price increases from T0 to T30. We now investigate this pattern.

We start by looking at how test price affects the average outcome for those who took the test. We investigate both average subjective diabetes risk and test outcome conditional on taking the test. If the test takers who remain are indeed the high-risk group, as predicted by the simple neoclassical intuition, these outcomes should differ significantly across treatments.

Figure 3 displays cross-treatment results. The left figure reports the mean value of subjective risk across treatments, together with the 90% confidence interval. The right figure displays the average diabetes test outcome in terms of blood glucose level across different treatments. For simplicity, we pool FPG and RPG outcomes and use GLU to denote the pooled outcome.[11] Despite a significant decline in the take-up rate as the price increases, both figures suggest *no* significant difference across treatments, either in terms of subjective diabetes risk or the actual outcome. This result contradicts the simple neoclassical prediction.

[Figure 3]

Table 2 reports formal regression results on how the price increase affects subjective risk and test outcome conditional on taking the test. The dependent variables are subjective risk (column (1)) and test outcome (column (2)). The key independent variables are the T10 and T30 treatment dummies. We control for six demographic variables—gender, age, education, marital status, household size, and monthly income—and 10 variables for health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. Consistent with Figure 3, we see no significant difference across treatments on these dependent variables after controlling for demographic information and health background.

[Table 2]

[11] The right figure reports results for the overall sample, including the fasting and non-fasting samples. For participants in the fasting state, GLU indicates fasting blood glucose level. For those in the non-fasting state, GLU indicates random blood glucose level. For simplicity, we pool the outcomes of the two tests here and add an indicator to control for whether it is a fasting blood sample in our analysis.

There are several possible explanations for why the mean test outcome does not change across treatments: Either the low- and high-risk groups select out of the test; both groups take the test and the medium-risk group selects out of the test; or individuals select out of the test independent of their disease risk. In the first case we expect to see a reduced dispersion of the test outcome, because test takers are more concentrated on the medium-risk level. In the second case the distribution should have more dispersion, while in the third case the dispersion should remain the same.

Figure 4 presents the distribution of test outcomes across treatments. The standard deviation of blood glucose level is 1.371 for T0; 1.484 for T10; and 0.733 for T30. As the price increases, there is a general trend of test outcome concentration toward the medium level, especially when the price increases to 30.

[Figure 4]

We perform two tests for the significance of the difference between these dispersion results, Levene's test and Brown-Forsythe's test, and report the *p* values in Table 3. In both tests, the difference in dispersion between T0 and T10 is not significantly different at the 10% level. However, when we compare T30 with T0 or T30 with T10, we can see that the variances are significantly different in most cases (the *p*-values are 0.075 and 0.121 for T30 vs. T0 and 0.021 and 0.023 for T30 vs. T10, respectively). When we pool T0 and T10 together to compare with T30, the difference in variance is also significant (the *p*-values are 0.049 and 0.072, respectively).

The above analysis suggests that as price increases, we observe no significant change in the average test outcome but significant reduction in the variance of test outcomes. These facts imply that both high- and low-risk subjects select out of the test in this process. These results are well identified, because prices are exogenously varied on random samples of participants.

[Table 3]

Finding 1: As the test price increases, both low- and high-risk groups select out of the test.

Note that due to the nature of the information avoidance phenomenon we study here, the relatively low take-up rate can lead to a potential power problem in the test; therefore, the results serve as suggestive evidence. However, together with other evidence presented later, they demonstrate a consistent pattern whereby both the high-risk and low-risk groups tend to select out of the test as the price increases.

3.2.2. Within-treatment results

Cross-treatment results suggest that both high-risk and low-risk groups select out of the sample as the price increases. Is there a similar pattern within treatments across different levels of subjective risk? A simple neoclassical model that assumes that individuals cannot take any action to treat the disease before being diagnosed predicts that the take-up rate is monotonically increasing in subjective risk—i.e., that high-risk individuals are more incentivized to take the test. We test this prediction first.

Figure 5 graphs the relationship between subjective risk and the take-up rate for T0, T10, and T30. Individuals are divided into five groups based on their subjective risk of diabetes: group 1 if subjective risk is 0; group 2 if subjective risk is between 0 and 0.2; and group 3, 4, or 5 if subjective risk is between 0.2 and 0.4; 0.4 and 0.6; and above 0.6, respectively. The x-axis indicates the group, while the y-axis is the average take-up rate within the group. The size of the data points is scaled to reflect the number of observations in each group. We can see that more than half of the subjects report no risk of diabetes, while the other groups have roughly comparable numbers of observations.

We can see that for treatments T0 and T10, the relationship is mostly monotonic; however, T30 shows a somewhat different pattern: The take-up rate is generally increasing when subjective risk increases, and the increasing tendency seems weaken as subjective risk increases.

[Figure 5]

Table 4 reports OLS regressions on the within-treatment pattern. The first three columns analyze patterns in the three treatments separately. The outcome variable is the dummy that indicates whether to take the diabetes test. The key explanatory variable is subjective risk and its square term. The control variable is the same as in Table 2. For T0 and T10, estimates of both subjective risk and its square term are not significantly different from zero, suggesting there is no concave relationship in T0 or T10. However, there is a concave relationship in T30: The estimate of subjective risk is significantly positive (0.82) and that of the square term is significantly negative (-1.10). These estimates suggest that the increase in test take-up with respect to subjective risk is decreasing in subjective risk. Results are robust when we add different category controls gradually. In particular, whether or not we control for risk-related factors (i.e., gender, age, and BMI) does not affect our results. For another robustness check, we construct five dummy variables that indicate different levels of subjective risk and report the results in Appendix Table A2. The pattern is largely consistent with the first three columns in Table

4, with estimates of the fifth dummy lower than that of the fourth dummy in T30, and the difference is marginally significant (p=0.108).

Another potential concern is that separate regressions allow extra flexibility that may misinterpret pure noise as a meaningful result. Also, due to the nature of the problem studied, the take-up rate is relatively low and there could be a power problem in the separate regressions. To address these concerns, we add two more columns in Table 4 by pooling observations from all three price experiments in a single regression. We add the interaction terms for T10 and T30 to subjective risk (column (4)) and its square term (column (5)). In column (4), we see that the take-up rate decreases when the price increases from 0 to 30 RMB across treatments, but there is no significant interaction effect between the price experiment dummies and subjective risk. In column (5), however, the estimated coefficient of "T30 × Subjective risk^2" is -1.57, significant at the 5% level, suggesting that compared with those in T0 and T10, the relationship in T30 is significantly more concave. The estimations are robust to using logistic regression. We report the logistic regressions of Tables 4-7 in Appendix Tables A3-A5.

[Table 4]

Finding 2: The increase in test take-up with respect to subjective risk is decreasing in subjective risk in T30. Compared with those in T0 and T10, the relationship in T30 is significantly more concave.

In general, the concave relationship in T30 is inconsistent with the simple neoclassical model, which predicts that the probability of taking the test should increase with subjective risk.

It is important to consider explanations other than information avoidance, because subjective beliefs within the treatment group are not exogenously varied. For instance, the high-risk group may have less health knowledge about the benefits of testing and subsequent medical treatment; they may also have higher compliance costs for undergoing treatment or are financially constrained from undergoing treatment if diagnosed with diabetes. All of these factors may contribute to test avoidance behavior.

To examine these alternatives, we directly test whether subjective risk is correlated with these variables. We measure health knowledge based on whether participants answer the knowledge questions on the survey correctly. These include subjective and objective knowledge of diabetes. We

measure treatment compliance cost based on questions about how difficult it would be to comply with diabetes treatment (Q58 and Q59). We measure financial status based on self-reported income (Q18 and Q19) and expenditure levels (Q12-14). For this analysis, we only use observations from the price experiment.

[Table 5]

Table 5 reports regression results. We can see that higher subjective risk is significantly correlated with better rather than worse subjective knowledge of diabetes. No other factors are significantly related to subjective risk. Therefore, the alternatives described above do not seem to explain the estimated non-monotonic pattern.

3.3. Disease experiment

This section reports results for the disease experiment, in which take-up decisions for diabetes and cancer tests are compared. We find that in the disease experiment, when there is neither monetary cost nor transaction cost (since the blood has already been drawn), about 14% and 11% of participants avoid the test for diabetes and cancer, respectively. This implies that even with no monetary cost and little transaction cost, test avoidance still exists, and this phenomenon is not predicted by the simple neoclassical model.

Figure 6 displays the take-up rate across treatments by different levels of subjective risk of having the corresponding disease. The y-axis is the take-up rate, and the x-axis denotes percentiles of subjective risk: group 1 if subjective risk is 0; group 2 if subjective risk is between 0 and 0.2; and group 3, 4, or 5 if subjective risk is between 0.2 and 0.4; 0.4 and 0.6; and above 0.6, respectively. The size of the data points is again scaled to reflect the number of observations in each group. We can see that more than half of the subjects report no risk of the corresponding disease, while the other groups have roughly comparable numbers of observations, except that the number of observations with risk higher than 0.6 is relatively small compared with other groups.

We can see that the take-up pattern is quite different for diabetes and cancer treatments. Both treatments start with a high take-up rate—around 0.8 to 0.9—when subjective risk is zero. This is because both tests are free and have no additional transaction cost, since the blood has already been drawn for other purposes. In the diabetes treatment, despite a slight drop when subjective risk is in the

middle range, the take-up rate generally increases with subjective risk and reaches about 1 when subjective risk is above 0.6. This pattern is consistent with group T0 in Figure 5 in the price experiment. In the cancer treatment, however, there is an obvious non-monotonic relationship in which the take-up rate is highest when subjective risk is between 0 and 0.2, then steadily drops as subjective risk increases.

[Figure 6]

Table 6 estimates the relationships in Figure 6. The first two columns estimate the relationships separately. Column (1) represents the diabetes treatment. Estimate of the effects of subjective risk suggest that for each 10% increase in subjective risk, the take-up rate will increase by 7.9%. The square term is negative, which suggests some evidence of a drop in the take-up rate for a high level of subjective risk, but is not significant. Column (2) represents the cancer treatment. We see a strong and significant non-monotonic effect in this case: Both the subjective risk and its square terms are significant, suggesting that the take-up rate first increases and then decreases with subjective risk. The estimated turning point is at a subjective risk level of 0.27, which is consistent with the pattern in Figure 6.

The next two columns pool observations for both diabetes and cancer treatments together, and introduce the interactions between the cancer treatment dummy and subjective risk (column (3)) and the square term of subjective risk (column (4)). In column (3), the estimated coefficient of "cancer × subjective risk" is -0.37, significant at the 5% level. This indicates that the tendency for the take-up rate to increase with subjective risk is significantly weaker in the cancer treatment than the diabetes treatment. In column (4), in which we allow for a nonlinear relationship, although the estimated coefficient of "cancer × subjective risk^2" is -0.71, none of the interaction terms is significant.

[Table 6]

Finding 3: There is a non-monotonic pattern between the subjective probability of having cancer and testing for cancer, but no evidence of a non-monotonic pattern for the free diabetes test. The tendency for the take-up rate to increase linearly with subjective risk is significantly weaker in the cancer treatment.

We further investigate whether subjects' evaluation of the severity of the disease would affect the take-up pattern. We add the variable "controllable," constructed from Q55 and Q69 in the survey. Q55 asks whether subjects believe that diabetes is curable, and Q69 asks whether subjects believe that

cancer can be controlled to some degree. The variable "controllable" takes the value 1 if the answer to Q55 is "Yes" and 0 otherwise in the diabetes treatment; it takes the value 1 if the answer to Q69 is "Yes" and 0 otherwise in the cancer treatment.

In the first two columns of Table 7, we again report estimations for the diabetes and cancer treatments separately. We add the variable controllable and interact controllable with subjective risk and its square term to determine whether the non-monotonic pattern differs by belief about the severity of the disease. Column (1) reports results for the diabetes treatment. The interaction terms show no significant effects, suggesting that the effect of subjective risk does not differ by whether individuals believe that diabetes is controllable or not. Column (2) reports results for the cancer treatment. In this case, the estimated coefficient of "subjective risk square" is -3.67 and that of "controllable × subjective risk^2" is 2.56. Both are significant, which indicates that there is a significant differential non-monotonic pattern: For those who believe that cancer is less controllable, the non-monotonic pattern is significantly stronger.

Columns (3) and (4) pool observations for the two treatments together. Column (3) introduces triple interaction terms between controllable, cancer treatment, and subjective risk (also interaction terms with the square of subjective risk in column (4)). In Column (3), the estimated coefficient of "cancer × subjective risk" is -0.80, significant at the 5% level. This suggests that the positive relationship between take-up and subjective risk is significantly weaker in the cancer treatment than the diabetes treatment for participants who believe the disease is less controllable. However, the triple interaction term "cancer × controllable × subjective risk" is not significant. Column (4) allows for the nonlinear relationship of subjective risk. As in column (3), for participants who believe the disease is less controllable, those in the cancer treatment have a significantly stronger non-monotonic relationship than those in the diabetes treatment (the estimated coefficient of "cancer × subjective risk^2" is -2.94, significant at the 1% level). The fact that the estimated coefficient of "cancer × controllable × subjective risk^2" is 3.50 and significant at the 10% level suggests that the difference between the cancer treatment and the diabetes treatment is also significantly larger for those who believe the disease is less controllable than for those who believe the disease is controllable.

[Table 7]

Finding 4: In the cancer treatment, those who believe the disease is less controllable

demonstrate a stronger non-monotonic pattern. For those with the less-controllable belief, the tendency for the take-up rate to increase linearly with subjective risk is significantly weaker, and the non-monotonic relationship is significantly stronger in the cancer treatment than in the diabetes treatment.

In general, the key message from the disease experiment is that when the disease is more severe and believed to be less controllable, the tendency to avoid taking the test is stronger.

As mentioned before, due to the low take-up rate in our sample, a single test may suffer from a potential power problem. However, the consistent evidence we obtain from (1) conducting both cross- and within-treatment analyses, (2) varying both price and disease type, and (3) the different outcome measures (e.g., subjective probabilities and test outcomes) gives us more confidence that the two primary empirical facts contain a meaningful message that is not due to pure noise—that is, first, that some people avoid taking the test when there is no cost of any form; and second, that both low- and high-risk individuals tend to avoid the test.

4. Theoretical Explanations

We derive two major findings from the empirical analysis: First, in the disease treatment when there is neither monetary cost nor transaction cost, some individuals still refuse to take the medical test; second, in the price treatment, as the test price increases, both low- and high-risk individuals tend to select out of the test. What theory can explain these two findings at the same time?

Three major groups of models help to explain the general phenomenon of information avoidance in the literature: the optimal expectations model (Brunnermeier and Parker, 2005; Oster et al., 2013); the attention model (Karlsson, Loewenstein, and Seppi, 2009; Golman and Loewenstein, 2015; Ganguly and Tasoff, 2016); and the curvature model (Caplin and Leahy, 2001; Caplin and Eliaz, 2003; Kőszegi, 2003; Eliaz and Spiegler, 2006). All three groups of models assume anticipatory utility—i.e., beliefs about the future affect current utility. However, the models differ in how anticipatory utility enters the utility function and whether subjects can hold biased beliefs. The optimal expectations model allows manipulation of beliefs, while the other two models maintain rational expectation assumptions. In this section, we illustrate the predictions of each group of models in our medical screening context in terms of these two empirical facts. Since all three classes of models introduce anticipatory utility to

the neoclassical model, we first describe the predictions of the neoclassical model as a benchmark case.

4.1. The neoclassical model

We first introduce the setup of the neoclassical model, which also applies to the other three anticipatory utility models more generally. We assume that there are two health states: the individual has the disease (diabetes or cancer) or does not. The individual has an initial belief about the probability of having the disease. At time 0, he chooses whether to learn the true state through medical testing. At time 1, he chooses a binary action about whether to take the treatment. If he takes the medical tests, he receives proper treatment based on the health status. Conditional on having the disease, taking the treatment is better than not doing anything. If he does not take the test, there are two possible assumptions in the literature: treatment is either possible or impossible without the test.

Under the assumption that proper treatment is possible only after formal medical test, testing is valuable, since the information may affect the subsequent treatment to cure the disease. However, the literature (e.g., Oster et al., 2013) sometimes also allows for alternative treatments, such as doing more exercise, changing dietary habits, or seeking other treatment, even without the test. This ancillary assumption is important for the predictions of all the models we discuss here. Therefore, we use the neoclassical model to illustrate the effects of this assumption. Specifically, we derive the predictions of the neoclassical model under two assumptions: treatment is and is not possible without the test.

We relegate the technical details and formal propositions to Appendix 3.1 and use Figure 7 to illustrate the relationship between risk and the value of testing in the neoclassical model under the two assumptions about treatment, assuming that the cost of the test is zero.[12] The value of testing on the y axis is the difference between the expected utility if tested and the expected utility if untested. Positive values of testing imply that individuals will choose to take the test. The dotted line represents a simple neoclassical model, which assumes that treatment of the disease can only be implemented after being formally diagnosed. The model predicts a 100% take-up rate if there is no cost of the test, and high-risk individuals should be more likely to take the test. The intuition is that the higher the risk, the larger the benefit from obtaining information and receiving the treatment. This is the instrumental value of taking the test.

[12] For the parameter values of the simulations illustrated in Figures 7 to 10, please refer to figure notes and Appendix 3.

The dashed line represents the neoclassical model under the assumption that treatment of the disease can be implemented without the test and the effect is as good as treatment after being diagnosed. The model also predicts 100% take-up with zero test cost, but it predicts a non-monotonic relationship: We observe that the benefit of testing first increases and then decreases as the risk becomes higher—i.e., the value of testing for the low- and high-risk groups is lower. The intuition is the following: When disease risk is greater than a threshold, the individual becomes more certain about having the disease, and hence might seek an equally effective alternative treatment even without the test. Therefore, the instrumental value of the testing is lower as risk becomes higher.

[Figure 7]

In general, Figure 7 suggests that the neoclassical model cannot explain information avoidance behavior when the test cost is zero, so we need to resort to models based on anticipatory utility. Comparing the dotted line and the dashed line, we can see that the ancillary assumption that treatment is possible without testing is important for generating the prediction that both high- and low-risk subjects tend to refuse the test. Notice that we can easily relax the assumption about perfectly effective treatment without the test to imperfect treatment—i.e., treatment is possible without testing but is less effective than treatment after formal diagnosis for untested individuals. The model prediction is similar (see Appendix 3.1): There is still a non-monotonic relationship. For simplicity, we adopt the simpler assumption—that the alternative treatment is equally effective—throughout our discussion. We will return to this point in Section 4.5.

4.2. The optimal expectations model

In this section we apply the theoretical model of Oster et al. (2013) to study take-up decisions in our setting (see Appendix 3.2 for detailed model setup and formal propositions). Their model is based on the optimal expectations model of Brunnermeier and Parker (2005). The idea of the model is that belief about future health status generates utility, which we call anticipatory utility. If individuals take the medical test, their beliefs about health status must update in a Bayesian way; i.e., their beliefs will be rational. They will also choose the correct treatment actions based on their health status. However, if individuals do not take the test, they are allowed to choose their beliefs based on the trade-off between the anticipation utility of feeling healthy today and the cost of wrong actions if they remain

ignorant. The influence of anticipation utility based on current beliefs creates the value of choosing overly optimistic beliefs, and this is only possible when one avoids taking the test.

Specifically, when individuals decide whether to take the test, they face the following trade-offs. The benefit of not testing is to hold biased beliefs, which generates high anticipation utility. The benefit of testing is to avoid the utility loss from the wrong state-matched action. If the former consideration outweighs the latter, individuals will choose not to test.

As discussed before, the ancillary assumption about whether treatment is possible without testing is crucial in terms of whether the predicted relationship between disease risk and the value of testing is non-monotonic. Figure 8 illustrates the model's prediction of the relationship between risk and the value of testing, again assuming that treatment is or is not possible without testing. For comparison, we also retain the dotted and dashed lines from the neoclassical model in Figure 7.

[Figure 8]

The asterisks line shows the optimal expectation model in which treatment is not possible without testing. The model predicts a monotonic relationship between risk and the value of testing. The intuition is that individuals face a tradeoff between the expected benefit of state-matched action (from testing) and the expected benefit of holding biased beliefs (from not testing), and these expected benefits are linear in disease risk in the model. If the benefit of testing is larger, we observe a monotonically increasing relationship. If the benefit of not testing is larger, we observe a monotonically decreasing relationship. In Figure 8, we illustrate the case in which the parameter values are such that the benefit of testing is larger; hence, there is an increasing relationship. Compared with the neoclassical model, in which treatment is not possible without testing (the dotted line), the possibility of holding biased beliefs about not testing and gaining higher anticipatory utility reduces the value of testing, and the decrease in the value of testing is larger as risk increases. However, this model still predicts a 100% take-up rate.[13]

The solid line shows the case in which equally effective treatment is assumed to be possible even without testing; this is the assumption Oster et al. (2013) use. The model predicts a non-monotonic relationship, in the sense that the benefit of testing first increases and then decreases as risk becomes

[13] We can also choose parameter values in which the benefit of not testing is larger, and hence there is a monotonically decreasing relationship—i.e. all individuals avoid the test even when it is free, and such tendency is higher for high-risk individuals. The case is similar for the attention-based model.

higher; the value of testing becomes negative for individuals with very high risk, and thus they refuse the test. The intuition is as follows: When risk is low, the benefit of not testing and holding biased beliefs is dominated by the benefit of testing to decide on the proper action. When risk is high, the negative total value of testing is mainly driven by two factors. First, the benefit of not testing and holding overly optimistic beliefs in terms of anticipatory utility is increasing in the risk level in this region. Second, the benefit of testing in terms of treatment benefit is decreasing in risk in this region. This is because when risk is higher than a certain threshold, individuals are assumed to take proper treatment action regardless of whether they test or not. As the test cost increases in the price treatment, this non-monotonic relationship can explain the fact that both low- and high-risk individuals avoid the test (see Figure A3 for details). In sum, Figure 8 illustrates how holding biased beliefs and allowing treatment without testing can generate information avoidance behavior even when the test is free. In addition, allowing treatment without testing generates a non-monotonic relationship, as in the case of the neoclassical model.[14] An optimal expectations model that allows treatment without testing can therefore explain two of our major findings.[15]

4.3. The attention model

We call the group of models that assume there is a discontinuous boost to utility when receiving information (e.g., Karlsson, Loewenstein, and Seppi, 2009; Golman and Loewenstein, 2015; Ganguly and Tasoff, 2016) the attention model. Different from the optimal expectations model, the attention model assumes rational expectation. In this section, we use Ganguly and Tasoff's (2016) model as a representative case, since it is applied in a health-screening setting.

The idea of Ganguly and Tasoff's (2016) model is that with the arrival of new information, individuals will experience a shock in anticipatory utility in the form of the difference between updated

[14] The assumption that proper treatment can be obtained even without formal diagnosis could be too strong. We also derive a model assuming imperfect treatment for untested individuals, and the predictions are similar to our main predictions as long as the imperfect treatment can still be somewhat effective (see Appendix 3.2, Propositions A7 and A8 for details).

[15] One caveat is that the predicted relationship in Figure 7 is between the total value of testing and the initial objective belief about the disease risk, but what we elicit in the field should be the updated subjective belief about the disease risk. Further, appendix 3 shows that the updated subjective belief is bimodal because the action space is binomial, but the empirical distribution of our subjective belief does not seem to be bimodal. Both facts suggest some potential inconsistency when linking the theoretical prediction with our empirical findings. Two ways can help alleviate these concerns: First, according to Appendix Propositions 1 in Oster et al. (2013), when action space is continuous, the subjective belief is also continuous, and it is an increasing function of true objective probability. Second, if we allow heterogeneity in the utility level of different health status, we can also observe a continuous subjective belief and a non-monotonic relationship between the value of testing and subjective beliefs. Thus, the value of testing should have the same relationship with respect to both the objective and subjective belief in a more general setting.

expected utility with new information and current utility (assumed to be independent of the disease risk in Ganguly and Tasoff).

[Figure 9]

Following Ganguly and Tasoff (2016), we formally set up the model and derive the prediction of the relationship between the expected value of testing and disease risk in Appendix 3.3. We simulate and illustrate the prediction in Figure 9.

The attention-based anticipatory utility model can generate information avoidance because of the following feature: individuals receive a discontinuous shock to utility when receiving information. The utility shock will be negative for potential bad news, and thus increases the benefit of not testing. In Ganguly and Tasoff (2016), treatment of disease is not possible without testing. We relax this assumption and show how the relaxed assumption affects information avoidance behavior as before.

The asterisks line in Figure 9 shows the attention model of Ganguly and Tasoff (2016) in which treatment is not possible without testing. Similar to the optimal expectations model, the model predicts a monotonic relationship between risk and the value of testing. The intuition is that both the expected benefit of state-matched action (from testing) and the expected benefit of avoiding negative attention-based utility shock (from not testing) are linear in disease risk. If the benefit of testing is larger/smaller, we observe a monotonically increasing/decreasing relationship. In Figure 9, we illustrate the case with parameter values in which the benefit of testing is larger, and thus the value of testing is increasing in disease risk. This implies a 100% take-up rate. Compared with the neoclassical model, in which treatment is not possible without testing (the dotted line), the negative utility shock reduces the value of testing, and this decrease in the value of testing is larger as risk increases.

The solid line represents the attention model from Ganguly and Tasoff (2016), in which we assume that treatment is possible without testing. In addition to predicting the information avoidance of high-risk individuals when the test is free, the model is able to generate the non-monotonic relationship whereby the value of testing is lower for both low- and high-risk individuals. The intuition is similar to the comparison in Figure 7: The benefit of testing first increases and then decreases as the risk becomes higher, since testing is less likely to affect subsequent treatment for high-risk individuals.

In sum, similar to the opitmal expectations model, Figure 9 demonstrates that an attention-based utility shock when assuming that treatment is possible without testing can explain our two major

empirical findings.

4.4. The curvature model

The third group of models assumes that information preferences depend on the curvature of the anticipatory utility function (e.g., Caplin and Leahy, 2001; Caplin and Eliaz, 2003; Kőszegi, 2003), which we call the curvature model. The model also assumes rational expectations, in the sense that belief is correct. In particular, we study the predictions of Kőszegi (2003) as the representative case, again because this model is applied to a health-screening setting.

The idea of Kőszegi's (2003) model is that belief about health status can affect utility. If belief is stochastic, the expected value is assumed to enter the utility function (see Appendix 3.4 for details).

[Figure 10]

The curvature model can generate information avoidance behavior due to the curvature of the anticipatory utility function. Given that expected beliefs about health status enter the utility function, the shape of the utility function determines the patient's preferences for information. If the anticipatory utility function is concave, individuals are information averse. The intuition is that a concave utility function implies diminishing marginal utility over beliefs about health status. Kőszegi (2003) does not specify whether treatment of disease is possible without testing. We specify two assumptions and show how these assumptions affect information avoidance behavior.

Assuming that individuals are information averse—i.e., the anticipatory utility function is concave—Figure 10 depicts the relationship between risk and the total value of testing in a curvature model under different assumptions. The dashed line shows the curvature model in which treatment of disease is not possible without testing. Kőszegi's (2003) model implies that the expected loss from not testing is monotonically increasing in risk: The higher the risk, the larger the loss if individuals avoid testing and take the wrong action. In the model, individuals face a tradeoff between the expected benefit of state-matched action (from testing) and the expected benefit of remaining uncertain due to information-averse utility (from not testing). The model then predicts that the value of testing is convex in disease risk, because the anticipatory utility is assumed to be concave. In Figure 10, we choose parameter values such that the value of testing first decreases and then increases with disease risk. In this case, the model also predicts information avoidance for low-risk individuals: When the benefit of

testing is low, the benefit of remaining uncertain based on the information-averse utility function dominates the benefit of state-matched action.[16]

The solid line represents the curvature model in which treatment of disease is possible without testing. Kőszegi's (2003) model implies that the expected loss from not testing is non-monotonic with respect to risk. When risk increases in the low-risk range, the loss first increases if individuals avoid the test and do not take treatment. When risk increases in the high-risk range, individuals who avoid the test will take treatment, and thus the loss will decrease as risk increases. The model therefore also predicts that the value of testing is lower for both low- and high-risk individuals.

Therefore, the information-averse utility function reduces the total value of testing. When we assume that treatment is possible without testing, the model predicts that both low- and high-risk individuals will avoid the test even when the test cost is zero. The curvature model that assumes that treatment is possible without testing can therefore explain our major findings.

4.5 Discussion

In sum, under the ancillary assumption that treatment is possible without the formal test, the three classes of anticipatory utility models can explain our findings based on different mechanisms. They can generate information avoidance through the following channels: In the optimal expectations model, holding biased beliefs increases the benefit of not testing and thus reduces the total value of testing. In the attention model, the attention-based anticipatory utility shock for potential bad news increases the benefit of not testing and thus reduces the total value of testing. In the curvature model, the information-averse utility function reduces the total value of testing. However, the finding that both low- and high-risk individuals are more likely to avoid the test requires an ancillary assumption common to all three classes of models: Individuals can take some form of alternative treatment action without being formally diagnosed.

Our empirical analysis reveals some important heterogeneity by price and disease type: The tendency for high-risk individuals to avoid the test is stronger when the price is higher (T30) and the disease is more severe (cancer). The models we examine can also explain these treatment effects in both the price and disease experiments. For the price treatment, Figures 7-10 depict the case in which

[16] We can also have parameter values in which the convex relationship is such that the value of testing always increases with disease risk. In this case, there is no test avoidance.

the test is free, but we can easily increase the test cost by moving the predicted relationship down in a parallel manner (see Figures A3-A5 for details). If the figures show a non-monotonic relationship, we can see that as the test price increases, both low- and high-risk individuals are more likely to avoid the test. In particular, the threshold risk for high-risk individuals to avoid the test (p_{high} in Appendix 3) is lower as the price increases, so we are more likely to observe the information avoidance phenomenon in high-risk individuals, given the fact that high-risk observations are scarce in the data. For the disease treatment, our simulation shows that as the disease becomes more severe, the threshold risk for high-risk individuals to avoid the test is also lower. Thus, in this case we are more likely to observe test avoidance in high-risk individuals.

The assumption that equally effective treatment is possible without the test seems extreme. However, as Appendices 3.1 and 3.2 show, the same conclusion holds even when we relax the assumption: We can assume that the alternative treatment is not as effective as the one received in the formal medical system after being diagnosed—but as long as its treatment effect is greater than a certain threshold, we get a similar non-monotonic relationship. On the survey, we did not observe clear patterns regarding how health behaviors are correlated with subjective risk, but there is weak evidence that individuals who avoid taking the diabetes test do more exercise and eat more vegetables as their subjective risk levels increase, especially when they are more knowledgeable about diabetes. We leave more focused research on this issue to future study.

5. Conclusion

This paper reports results from a randomized field experiment in rural China to investigate whether individuals have a tendency to avoid medical tests due to information avoidance. We randomly assigned individuals to different treatments that varied the price of a diabetes test and different treatments to vary the type of disease being tested (diabetes or cancer). We have two major findings. First, some individuals avoid the test when there is neither monetary nor transaction cost in the disease treatment. Second, in the price treatment, we observe that as the test price increases, the average test outcome remains the same, but the dispersion of the outcome decreases, which indicates that both low- and high-risk individuals select out of the test as the price increases. This phenomenon has not previously been revealed. We also find interesting heterogeneity: The pattern by which high-risk

individuals avoid the test is more salient when the test price is higher and when the disease is more severe.

We apply three groups of models of information avoidance to our setting and derive their predictions on the relationship between the probability of having the disease and the net value of testing. We find that with the additional ancillary assumption that somewhat effective treatment is possible without the test, all three models can simultaneously explain the two major empirical findings. The models can also predict the heterogeneity across test price and disease type in our empirical findings.

How the tendency to avoid information varies across the probability of having the disease has important policy implications. When policymakers mainly consider the instrumental value of testing (in a simple neoclassical model) or the disease has negative externality, the test is more valuable for high-risk individuals. However, according to our empirical results, some high-risk individuals are less likely to take the test. In this case, proper interventions that target the high-risk group yield higher welfare gains than traditionally assumed. Also, new policies that target the group that attaches greater weight to anticipatory utility can be more effective than traditional policies.

References

Allcott, Hunt, and Dmitry Taubinsky (2015). "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market." *American Economic Review*, Vol 105, No. 8 (August), pages 2501-2538.

Beshears, John, James Choi, Christopher Clayton, Christopher Harris, David Laibson, Brigitte Madrian. 2017. "Optimal Illiquidity" Working paper.

Bénabou, Roland, and J. Tirole, "Self Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117 (2002), 871–915.

Brunnermeier, Markus K., and J. A. Parker. "Optimal Expectations." *American Economic Review* 95.4(2005):1092-1118.

Caplan, L. S. "Patient Delay in Seeking Help for Potential Breast Cancer. " *Public Health Reviews* 23.3(1995):263-74.

Caplin, Andrew, and J. Leahy. "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics* 116.1(2001):55-79.

Eliaz, Kfir, and R. Spiegler. "Can anticipatory feelings explain anomalous choices of information sources?" *Games & Economic Behavior* 56.1(2006):87-104.

Epstein, Larry G. "Living with Risk." *The Review of Economic Studies* 75.4(2008): 1121-1141

Golman, Russell, D. Hagmann, and G. Loewenstein. "Information Avoidance." *Journal of Economic Literature* 55(2017):96-135.

Ganguly, Ananda R, and J. Tasoff. "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future." *Management Science* 63(2016):4037-4060.

Golman, Russell and G. Loewenstein. "An Information-Gap Framework for Capturing Preferences About Uncertainty." *Proceedings of the 15th Conference on Theoretical Aspects of Rationality and Knowledge* (2015)

Hariri S, Yoon PW, Qureshi N, Valdez R, Scheuner MT, Khoury MJ: Family history of type 2 diabetes: a population-based screening tool for prevention? Genet Med 8:102–108, 2006

Holt, Charles A., and S. K. Laury. "Risk Aversion and Incentive Effects." *American Economic Review* 92.5(2002):1644-1655.

International Diabetes Federation 2015. Brussels: International Diabetes Federation

Karlsson, Niklas, G. Loewenstein, and D. Seppi. "The ostrich effect: Selective attention to information." *Journal of Risk and Uncertainty* 38.2 (2009): 95-115.

Kőszegi, Botond. "Health anxiety and patient behavior." *Journal of Health Economics* 22.6(2003):1073-1084.

Kőszegi, Botond and Matthew Rabin. "Reference-dependent consumption plans." *The American Economic Review*, 99.3(2009):909-936.

Kőszegi, Botond, "Utility from Anticipation and Personal Equilibrium," *Economic Theory*, 44 (2010), 415–444.

Lerman, C, et al. "Genetic Testing in Families with Hereditary Nonpolyposis Colon Cancer. " *JAMA* 281.17(1999):1618.

Liebman, J. and R. Zeckhauser, "Simple Humans, Complex Insurance, Subtle Subsidies," Working Paper 14330, NBER 2008.

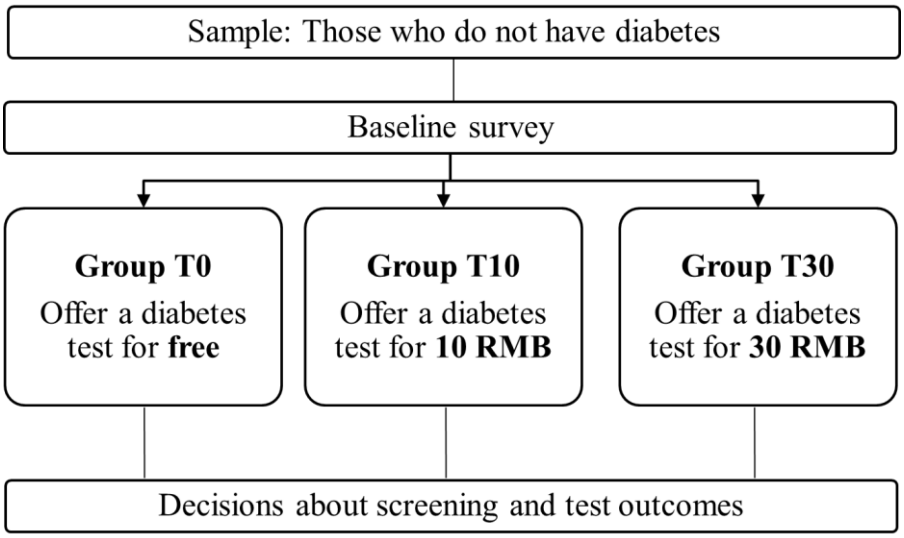
Lyter, D. W., et al. "The HIV antibody test: why gay and bisexual men want or do not want to know their results. " *Public Health Reports*102.5(1987):468.

Meechan, G, J. Collins, and K. Petrie. " Delay in Seeking Medical Care for Self-Detected Breast

Symptoms in New Zealand Women." *New Zealand Medical Journal* 115 (1166):U257.
Miller, S. M. "Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat." *Journal of Personality & Social Psychology* 52.2(1987):345-53.
Nosarti, C., Crayford, T., Roberts, J. V., Elias, E., McKenzie, K., & David, A. S. Delay in presentation of symptomatic referrals to a breast clinic: patient and system factors. *British journal of cancer*, 82.3 (2000): 742.
Okeke, E. N., C. A. Adepiti, and K. O. Ajenifuja. "What is the price of prevention? New evidence from a field experiment." *Journal of Health Economics* 32.1(2013):207.
Oster, Emily, I. Shoulson, and E. R. Dorsey. "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease." *American Economic Review* 103.2(2013):804-830.
Persoskie, Alexander, Rebecca A. Ferrer, and William M. P. Klein. 2014. "Association of Cancer Worry and Perceived Risk with Doctor Avoidance: An Analysis of Information Avoidance in a Nationally Representative US Sample." *Journal of Behavioral Medicine* 37 (5): 977–87.
Sullivan, Patrick S., Amy Lansky, and Amy Drake. 2004. "Failure to Return for HIV Test Results among Persons at High Risk for HIV Infection: Results from a Multistate Interview Project." *JAIDS Journal of Acquired Immune Deficiency Syndromes 35 (5): 511–18.*
Thornton, R. L. "The Demand for, and Impact of, Learning HIV Status." *American Economic Review* 98.5(2008):1829-1863.
World Health Organization (2016), *Global Report on Diabetes*. Geneva, 2016.
World Health Organization (2017), "The top 10 causes of death Fact sheet". *World Health Organization*. Jan 2017.
Yang, W., et al. "Prevalence of diabetes among men and women in China." *New England Journal of Medicine* 362.25(2010):2425.

Figure 1. Experimental Design

Panel A: Price experiment



Panel B: Disease experiment

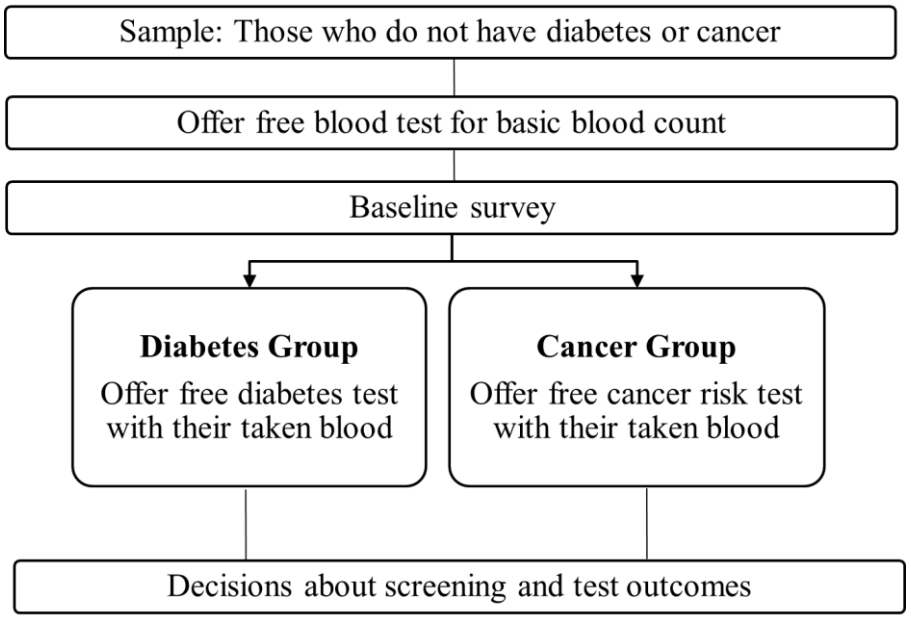
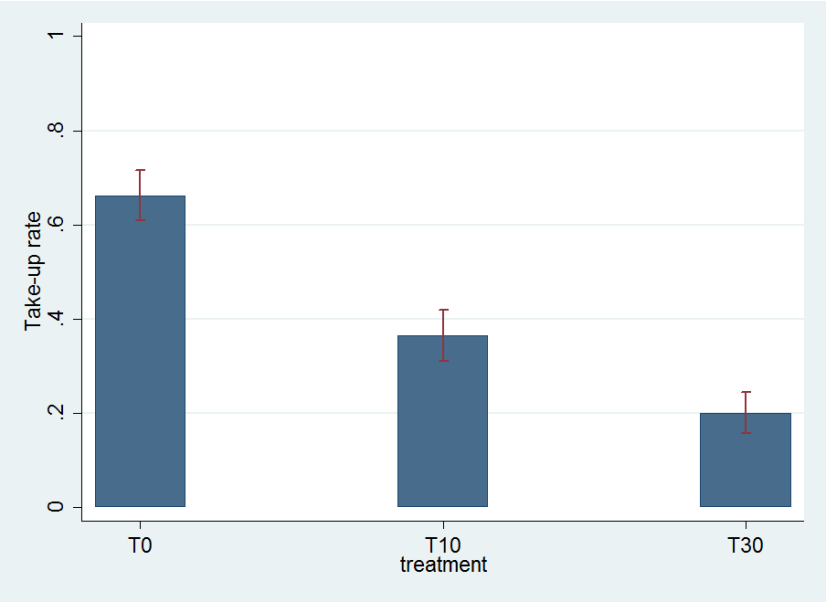
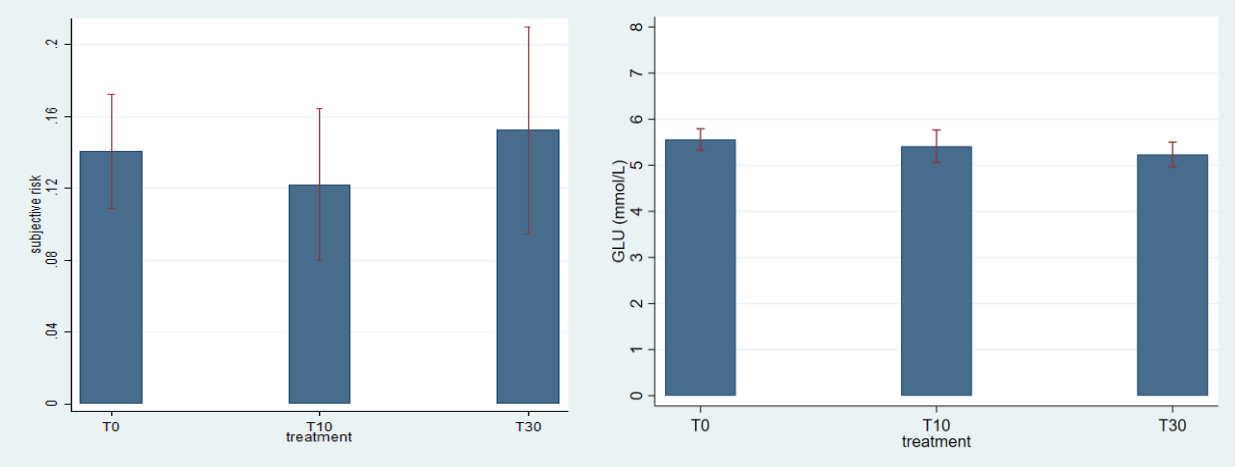


Figure 2. Take-up Rate across Treatments



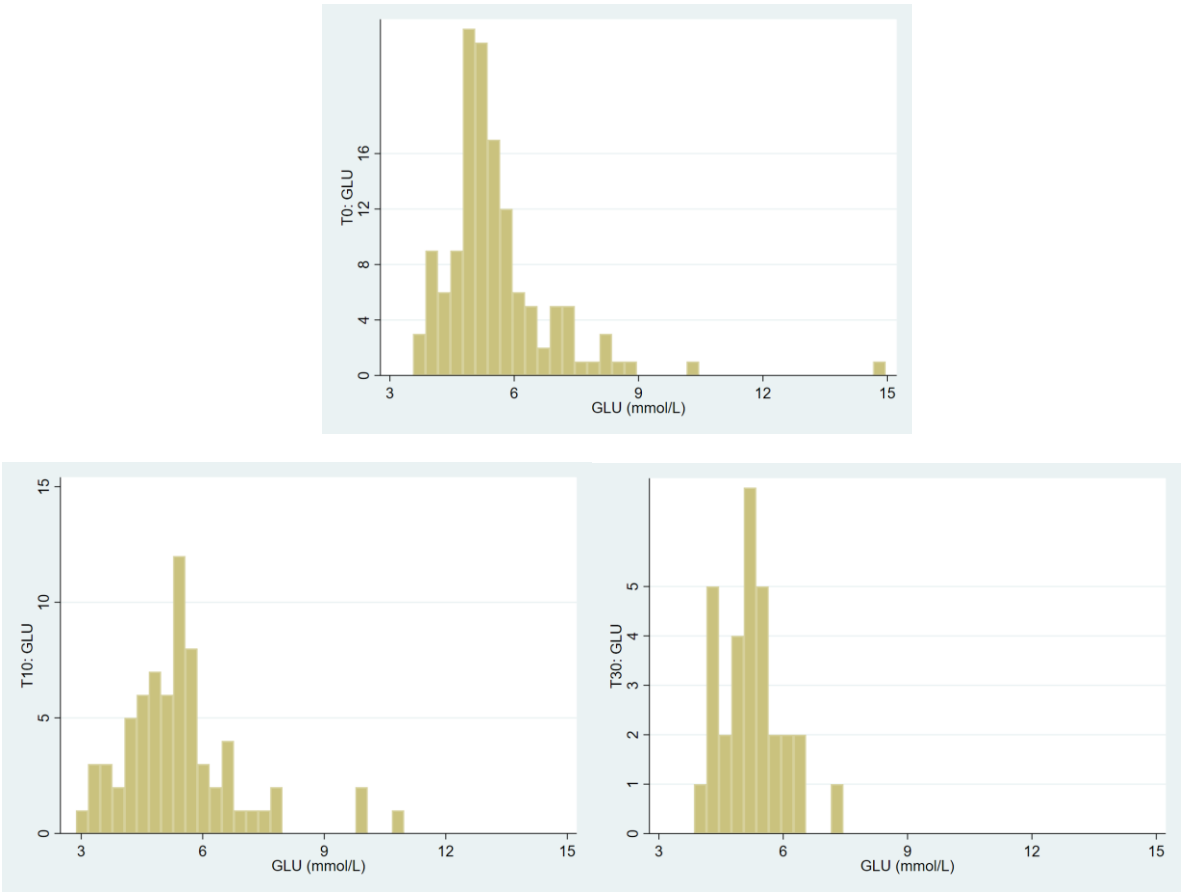
Note: This figure compares subjects' take-up rates of the diabetes test across different treatments with 90% confidence intervals.

Figure 3. Risk and Test Outcome Conditional on Taking the Test



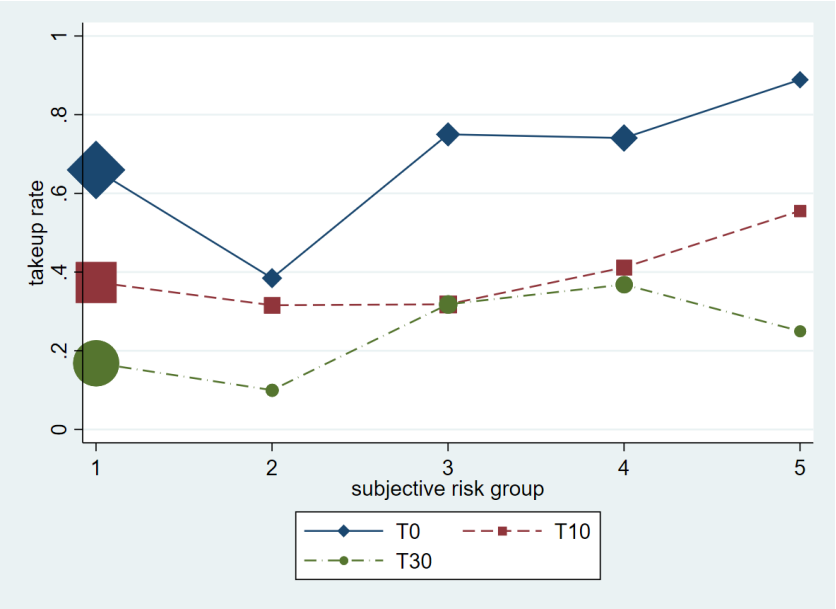
Note: The left figure displays the average subjective risk across different treatments, with a 90% confidence interval conditional on taking the test. Subjective risk is the chance that individuals think of themselves as having diabetes. The right figure displays the average blood glucose level (GLU) of the two diabetes tests across different treatments conditional on taking the test.

Figure 4. Distribution of Test Outcomes across Treatments



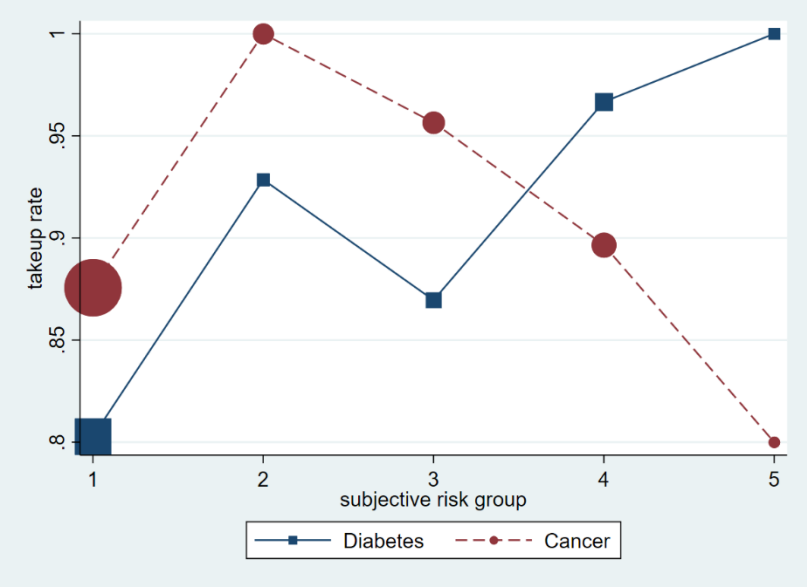
Note: This figure displays the frequency distributions of blood glucose level (GLU) of the two diabetes tests across different treatments conditional on taking the test.

Figure 5. Relationship between Subjective Risk and the Take-up Decision



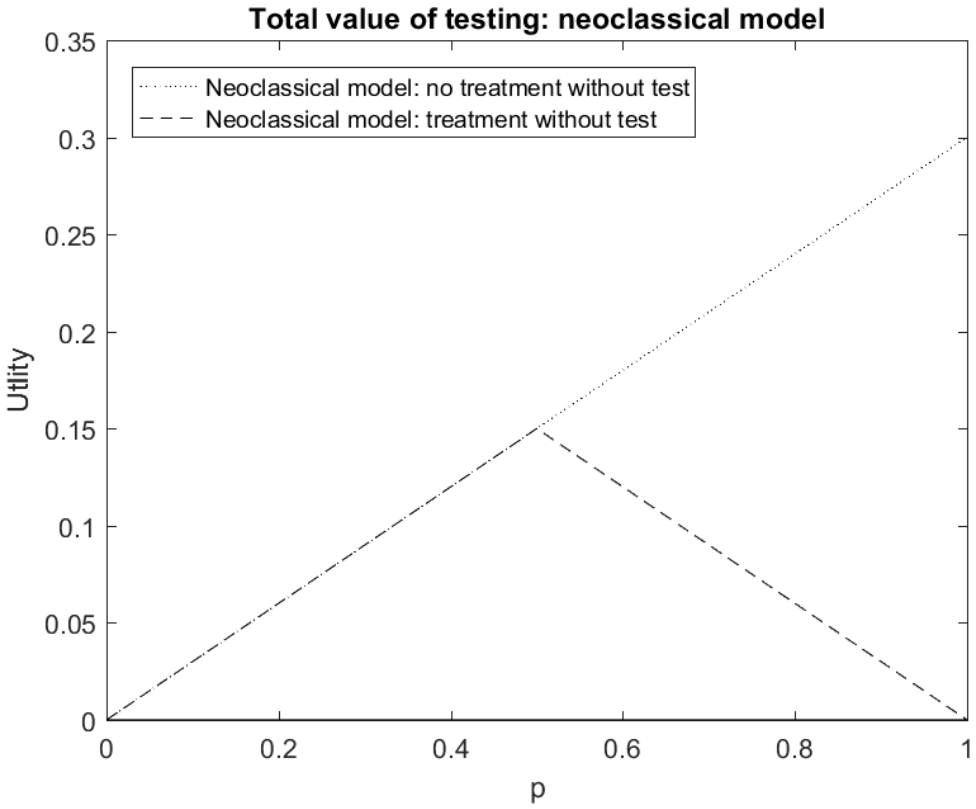
Note: Individuals are divided into 5 groups based on their subjective risk of diabetes: group 1 if subjective risk is 0; group 2 if subjective risk is between 0 and 0.2; and group 3, 4, or 5 if subjective risk is between 0.2 and 0.4, 0.4 and 0.6, or above 0.6, respectively. We scale the size of the data points to the number of observations in each group.

Figure 6. Take-up Rate across Treatments by Percentiles of Subjective Risk of the Disease



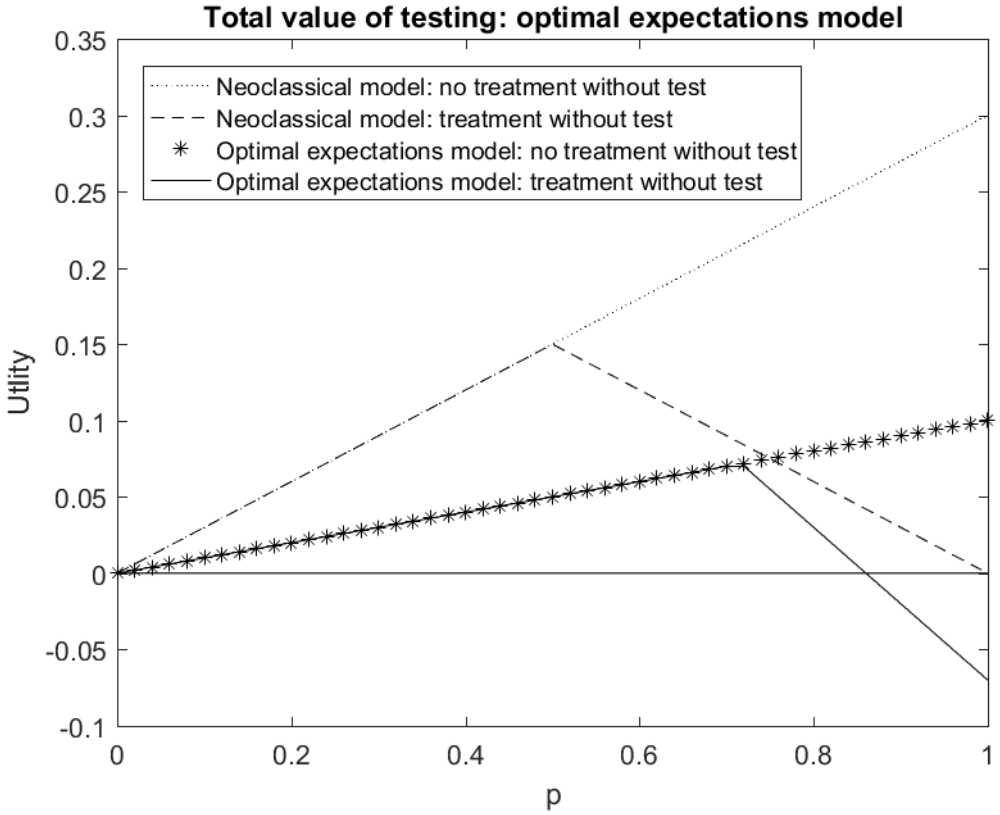
Note: Individuals are divided into 5 groups based on their subjective risk of the corresponding disease: group 1 if subjective risk is 0; group 2 if subjective risk is between 0 and 0.2; and group 3, 4, or 5 if subjective risk is between 0.2 and 0.4, 0.4 and 0.6, or above 0.6, respectively. For individuals in the diabetes treatment, subjective risk is defined as the chance that they believe they will develop diabetes. For individuals in the cancer treatment, it is defined as the chance that they believe they will develop cancer. We scale the size of the data points to the number of observations in each group.

Figure 7. Total Value of Testing in the Neoclassical Model



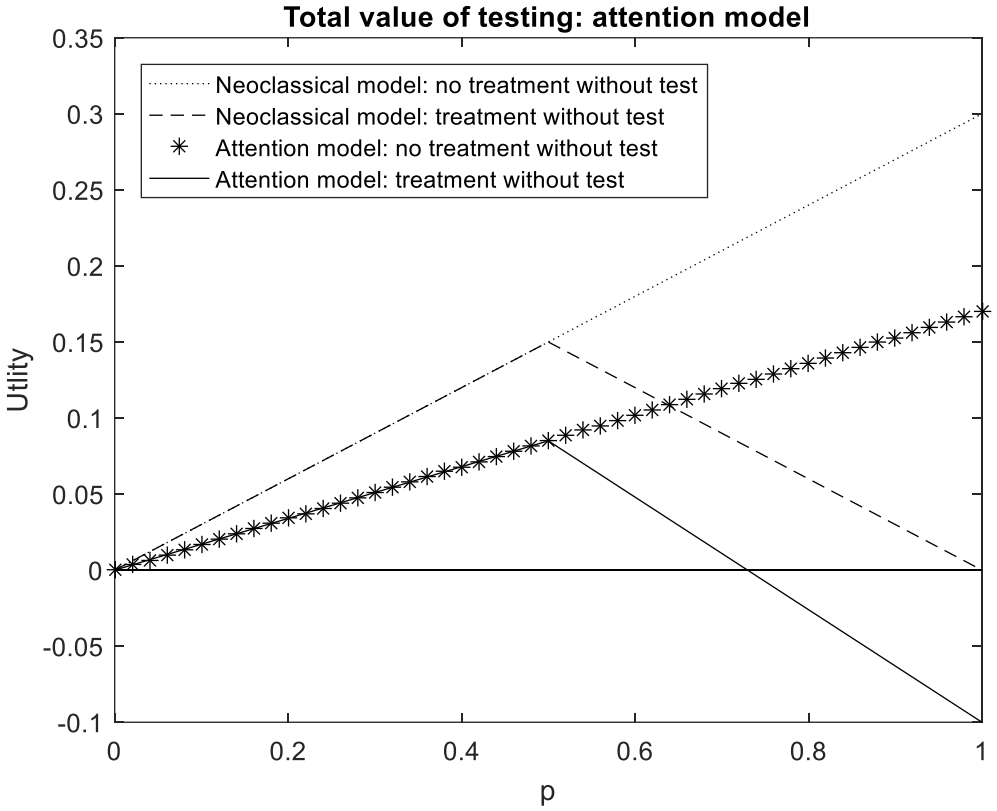
Note: This figure shows the total value of testing when the test is free. The horizontal axis is probability of having the disease. The vertical axis is the total value of testing, which equals the benefit of testing minus the benefit of not testing. We choose $\Omega = 0.3$, $\Phi = 0.3$ and $v = 0$ in the neoclassical model.

Figure 8. Total Value of Testing in Optimal Expectations Model



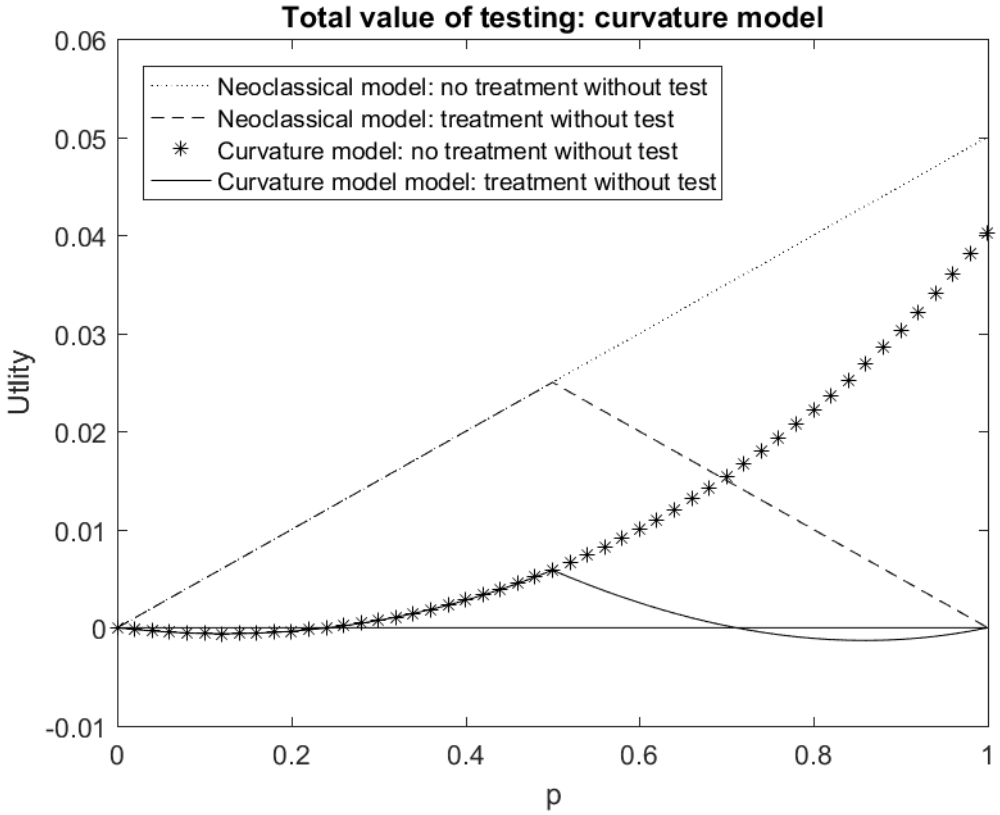
Note: This figure shows the total value of testing when the test is free. The horizontal axis is probability of having the disease. The vertical axis is the total value of testing, which equals the benefit of testing minus the benefit of not testing. We choose $\delta = 0.2$, $\Omega = 0.3$, $\Phi = 0.3$, and $v = 0$ in the optimal expectations model.

Figure 9. Total Value of Testing in the Attention Model



Note: This figure shows the total value of testing when the test is free. The horizontal axis is probability of having the disease. The vertical axis is the total value of testing, which equals the benefit of testing minus the benefit of not testing. We choose $a(\mu_0, \mu_1)$=0.1, $u_p = 1, \Omega = 0.3$, $\Phi = 0.3$ in the attention model.

Figure 10. Total Value of Testing in the Curvature Model



Note: This figure shows the total value of testing when the test is free. The horizontal axis is probability of having the disease. The vertical axis is the total value of testing, which equals the benefit of testing minus the benefit of not testing. In the curvature model, we choose curvature with information avoidance $\alpha = 0.8$, $s_l = 1$, $s_h = 2$. When individuals do not take the test, $l(s, t^*) = \begin{cases} 0.05 & if \ \ s = s_l \\ 0 & if \ \ s = s_h \end{cases}$.

Table 1. Summary Statistics and Randomization Check

	T0	T10	T30	Diabetes	Cancer
Panel A. Screening					
Take-up rate of the test	0.66	0.37	0.20	0.86	0.89
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Panel B. Demographics					
Gender (male)	0.37	0.38	0.43	0.45	0.39
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Age	53.29	51.57	52.59	52.61	52.05
	(0.45)	(0.47)	(0.48)	(0.47)	(0.43)
Education years	7.28	7.90	6.67	7.04	7.09
	(0.21)	(0.19)	(0.22)	(0.21)	(0.19)
Marriage Status	0.94	0.98	0.94	0.90	0.93
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Household Size	3.18	3.54	3.27	3.29	3.35
	(0.09)	(0.10)	(0.09)	(0.09)	(0.09)
Whether monthly income is	0.48	0.55	0.48	0.54	0.45**
larger than 1,000 RMB	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Panel C. Health Conditions and Behaviors					
Height (cm)	159.46	160.28	160.17	160.91	160.43
	(0.52)	(0.54)	(0.54)	(0.55)	(0.47)
Weight (kilogram)	67.13	65.57	68.39	67.36	66.79
	(0.74)	(0.80)	(0.82)	(0.78)	(0.66)
BMI ratio	26.44	25.49	26.63	25.98	25.96
	(0.29)	(0.27)	(0.28)	(0.27)	(0.24)
Smoking (percentage)	0.30	0.31	0.36	0.36	0.34
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Drinking (percentage)	0.35	0.31	0.33	0.42	0.41
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Sleeping hours	7.73	7.84	7.51	7.74	7.85
	(0.11)	(0.10)	(0.12)	(0.10)	(0.09)
Exercise frequency	2.63	2.74	2.80	2.76	2.75
	(0.09)	(0.09)	(0.09)	(0.09)	(0.08)
Subjective knowledge of					
diabetes	0.31	0.30	0.28	0.30	0.29
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Objective knowledge of					
diabetes	0.47	0.48	0.46	0.46	0.44
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Ability to follow treatment	1.44	1.14	1.40	0.84	0.83
	(0.43)	(0.31)	(0.58)	(0.01)	(0.01)
	0.12	0.11	0.10	0.13	0.10

Subjective assessment of disease risk	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Panel D. Preference Coefficients					
Risk aversion	3.13	3.30	3.17	2.87	3.18
	(0.15)	(0.15)	(0.15)	(0.15)	(0.14)
Loss aversion	2.46	2.65	2.68	2.47	2.58
	(0.17)	(0.17)	(0.17)	(0.17)	(0.15)
Patience (includes present bias)	3.28	3.15	3.50	3.28	3.51
	(0.18)	(0.18)	(0.17)	(0.17)	(0.16)
Patience (not including present bias)	3.16	2.98	3.47	3.22	3.49
	(0.18)	(0.18)	(0.18)	(0.17)	(0.16)
Monitoring	0.13	0.13	0.13	0.13	0.13
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Neuroticism	2.52	2.62	2.58	2.63	2.69
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
Openness	4.67	4.57	4.70	4.52	4.52
	(0.06)	(0.06)	(0.05)	(0.07)	(0.06)
Observations	219	216	229	255	276

Note: We use * on the T30 variable to indicate whether the variables in T0, T10, and T30 are significantly different in the multivariate test, and whether variables in the cancer treatment are significantly different from those in the diabetes treatment.

Table 2. Subjective Risk and Test Outcomes Conditional on Taking the Test

	(1)	(2)
	Subjective risk	GLU
T10	-0.03	-0.04
	(0.03)	(0.22)
T30	-0.02	-0.31
	(0.04)	(0.21)
Constant	-0.12	1.96
	(1.74)	(12.12)
Demographics (6)	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes
Observations	254	231
R-squared	0.09	0.18
F-statistics: T10=T30	0.164	0.87

Note: Regressions in the table show average subjective risk and test outcomes across treatments conditional on taking the test. We control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and 10 variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table 3. *p*-values of Equal-variance Tests on Medical Test Outcomes across Treatment

Tests	Levene	Brown-Forsythe
T0 vs T10	0.411	0.314
T0 vs T30	0.075	0.121
T10 vs T30	0.021	0.023
(T0+T10) vs T30	0.049	0.072

Note: This table reports the *p*-values of equal-variance tests on medical test outcomes across treatment. Both Levene's test and Brown-Forsythe's test are reported.

Table 4. Testing the Effect of Subjective Risk on Take-up Decision: Price Experiment

	(1)	(2)	(3)	(4)	(5)
	T0	T10	T30	Full Sample	Full Sample
Subjective risk	-0.02	-0.39	0.82*	0.35***	0.02
	(0.41)	(0.46)	(0.43)	(0.12)	(0.38)
Subjective risk^2	0.55	0.94	-1.10**		0.55
	(0.57)	(0.61)	(0.54)		(0.53)
T10				-0.26***	-0.25***
				(0.06)	(0.06)
T30				-0.44***	-0.46***
				(0.05)	(0.05)
T10 × Subjective risk				-0.22	-0.47
				(0.22)	(0.59)
T10 × Subjective risk^2					0.40
					(0.82)
T30 × Subjective risk				-0.20	0.74
				(0.20)	(0.58)
T30 × Subjective risk^2					-1.57**
					(0.78)
Constant	8.77**	3.62	-2.28	2.67	2.71
	(4.04)	(4.97)	(3.76)	(2.31)	(2.31)
Demographics (6)	Yes	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes	Yes
Observations	204	197	209	610	610
R-squared	0.10	0.09	0.12	0.20	0.20

Note: The regressions test the effects of subjective risk on take-up decisions. Results are not affected by adding different categories of controls gradually. Columns (1)-(3) estimate the effect separately in T0, T10, and T30, respectively. Columns (4) and (5) pool observations from T0, T10, and T30 together to test the cross-treatment difference. We control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and 10 variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table 5. Alternative Explanations

	(1)
	Subjective risk
Subjective Knowledge	0.10**
	(0.04)
Objective Knowledge	0.05
	(0.04)
Treatment Compliance Cost	-0.01
	(0.01)
Income Level	-0.01
	(0.01)
Expenditure Level	0.02
	(0.02)
Constant	0.15
	(0.12)
Demographics (6)	Yes
Observations	620
R-squared	0.05

Note: The regression in the table shows the effects of subjective and objective knowledge of diabetes, treatment compliance cost, income, and expenditure level on subjective risk. We control for six demographic variables: gender, age, education, marriage, household size, and monthly income. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table 6. Testing the Effect of Subjective Risk on Take-up Decision: Disease Experiment

	(1) Diabetes treatment	(2) Cancer treatment	(3) Full sample	(4) Full sample
Subjective risk	0.79**	0.74*	0.42***	0.79**
	(0.33)	(0.39)	(0.09)	(0.33)
Subjective risk square	-0.69	-1.39*		-0.69
	(0.52)	(0.78)		(0.52)
Cancer			-1.57	-1.50
			(4.33)	(4.28)
Cancer × Subjective risk			-0.37**	-0.05
			(0.16)	(0.51)
Cancer × Subjective risk^2				-0.71
				(0.94)
Constant	1.89	0.39	1.82	1.89
	(2.08)	(3.74)	(2.10)	(2.07)
Demographics (6)	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes
Observations	211	239	450	450
R-squared	0.10	0.10	0.09	0.10

Note: This table tests the effect of subjective risk on take-up decisions in disease experiment. Results are not affected by adding different categories of controls gradually. Columns (1) and (2) estimate the effect separately in the diabetes and cancer treatments. Columns (3) and (4) pool observations from the two treatments together to test the cross-treatment difference. In the first two columns, we control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and 10 variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. In the last two columns, we additionally control for the interaction terms of Cancer treatment with six demographic variables and 10 variables on health conditions and behaviors, because the effects of these variables on the two diseases can be very different. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table 7. Heterogeneous Effect in Disease Experiment

	(1) Diabetes treatment	(2) Cancer treatment	(3) Full sample	(4) Full sample
Subjective risk	0.77**	1.48***	0.37***	0.77**
	(0.38)	(0.43)	(0.11)	(0.38)
Subjective risk^2	-0.73	-3.67***		-0.73
	(0.58)	(0.77)		(0.58)
Cancer			-0.69	-0.39
			(4.37)	(4.38)
Cancer × Subjective risk			-0.80**	0.71
			(0.37)	(0.57)
Cancer × Subjective risk^2				-2.94***
				(0.97)
Controllable	-0.23**	-0.07	-0.22**	-0.23**
	(0.11)	(0.06)	(0.10)	(0.11)
Controllable × Subjective risk	0.88	-0.76	0.35	0.88
	(0.93)	(0.60)	(0.26)	(0.93)
Controllable × Subjective risk^2	-0.94	2.56**		-0.94
	(1.41)	(1.10)		(1.40)
Cancer × Controllable			0.13	0.16
			(0.12)	(0.13)
Cancer × Controllable × Subjective risk			0.26	-1.64
			(0.45)	(1.11)
Cancer × Controllable × Subjective risk^2				3.50*
				(1.79)
Constant	0.92	0.53	0.83	0.92
	(2.08)	(3.84)	(2.07)	(2.07)
Demographics (6)	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes
Observations	176	207	383	383
R-squared	0.17	0.16	0.15	0.17

Note: This table tests the heterogeneous effect of subjective risk on take-up decisions in the disease experiment, depending on whether individuals believe the disease to be controllable or not. Results are not affected by adding different categories of controls gradually. Columns (1) and (2) estimate the effect separately in the diabetes and cancer treatments. Columns (3) and (4) pool observations from the two treatments together to test the cross-treatment difference. In the first two columns, we control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and 10 variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. In the last two columns, we additionally control for the interaction terms of Cancer treatment with six demographic variables and 10 variables on health conditions and behaviors, because the effects of these variables on the two diseases can be very different. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

(For Online Publication Only)

Appendix 1. Introduction on Diabetes Test

1. Random (also called Casual) Plasma Glucose Test
   a) Procedure: This test is a blood check at any time of the day when you have severe diabetes symptoms.
   b) Criteria: Diabetes is diagnosed at blood glucose of greater than or equal to 200 mg/dl (11mmol/l).
   c) The test is not so good at diagnosing diabetes in people with mildly elevated blood sugar levels but it is good for those that have a high blood sugar level and may need treatment more urgently. (http://www.diabetes.co.uk)
   d) Price: $25 according to CVS
   e) Source: American Diabetes Association unless otherwise stated
2. Fasting Plasma Glucose (FPG)：
   a) Procedure: This test checks your fasting blood glucose levels. Fasting means after not having anything to eat or drink (except water) for at least 8 hours before the test. This test is usually done first thing in the morning, before breakfast.
   b) Criteria: Diabetes is diagnosed at fasting blood glucose of greater than or equal to 126 mg/dl (7 mmol/l).
   c) It is often the first test done to check for prediabetes and diabetes.
   d) Price: The cost of a fasting plasma glucose will usually cost $5 to $40. (howmuchisit.org)
   e) Source: American Diabetes Association unless otherwise stated
3. Oral Glucose Tolerance Test (also called the OGTT)
   a) Procedure: The OGTT is a two-hour test that checks your blood glucose levels before and 2 hours after you drink a special sweet drink. It tells the doctor how your body processes glucose.
   b) Criteria: Diabetes is diagnosed at 2 hour blood glucose of greater than or equal to 200 mg/dl (11mmol/l)
   c) An OGTT is the only means of identifying people with Impaired Glucose Tolerance;
      An OGTT is frequently needed to confirm or exclude an abnormality of glucose tolerance in asymptomatic people. (WHO)
   d) Price: Seem to be around $140
   e) Source: American Diabetes Association unless otherwise stated

4. A1C test

a) Procedure: To measure a person's HbA1c level, a blood sample is taken from the patient's arm, and used to produce a reading. (http://www.diabetes.co.uk)

   The A1C test measures your average blood glucose for the past 2 to 3 months. The advantages of being diagnosed this way are that you don't have to fast or drink anything.

b) Criteria: Diabetes is diagnosed at an A1C of greater than or equal to 6.5%.

c) The HbA1c test, also known as the haemoglobin A1c or glycated haemoglobin test, is an important blood test that gives a good indication of how well your diabetes is being controlled.

   Together with the fasting plasma glucose test, the HbA1c test is one of the main ways in which type 2 diabetes is diagnosed.

   HbA1c tests are not the primary diagnostic test for type 1 diabetes but may sometimes be used together with other tests. (http://www.diabetes.co.uk)

d) Price: Around $35 (Walgreens and CVS pharmacy)

e) Source: American Diabetes Association unless otherwise stated

5. CEA (carcinoembryonic antigen) is a blood test commonly used to follow patients with known cancers. CEA is a glycoprotein (sugar protein) present in embryonic tissues and in extracts from normal colonic washings. The test should not be used as a cancer screening test of asymptomatic individuals. Although CEA levels are often elevated in patients with gastrointestinal malignancies (colon, pancreas, etc), patients with confirmed cancers frequently have normal levels (in the range of healthy individuals). Elevations in CEA levels may occur in patients without cancer. For example, elevated CEA levels may be observed in smokers as well as patients with a variety of non-malignant diseases. Therefore, levels, regardless of their values cannot be used as a diagnostic test for cancer. The greatest value of these tests is in detecting recurrence of malignancy (cancer) after treatment of the tumor. Finally, remember that the normal values of these tests differ from lab to lab. Values obtained from one lab cannot be directly compared to those obtained from another lab.

Source: The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins (http://pathology.jhu.edu/pc/bloodmarkers.php)

Appendix 2. Surveys

Village:_____ Interviewee ID:_____ Interviewer:_____ Date:_____ Time: _____

0. Is the questionnaire completed?

0.1 Why the questionnaire is incomplete:

1) The interviewee refused to answer. 2) The interviewee was diagnosed with diabetes. 3) Other interviewee-related reasons 4) Mistakes of the interviewer

1. Have you been diagnosed with the following diseases? _____

1) Hypertension 2) Diabetes 3) Coronary disease 4) Stroke 5) Tumor 6) Psychiosis 7) Others_____

8) None

(The survey only continues when diabetes is not mentioned in this question.)

Part One: General Information

2. Name_____

3. Address_____

4. Phone Number_____

5. Gender: 1) Male 2) Female

6. Date of Birth：_____Year _____Month _____Day *Solar Calendar

6.1 Type of Calendar: 1) Solar calendar 2) Lunar calendar

7. Highest achieved education:

1) Primary school or lower 2) Junior high school 3) Senior high school 4) Secondary vocational school 5) Junior college or bachelor degree 6) Master or higher

8. How many years have you been in school? _____

9. Besides your formal school education mentioned above, what types of professional education have you received?

_____

10. Marital Status: 1) Single 2) Married 3) Divorced 4) Widowed 5) Separated

11. How many people in your family? _____ Number of kids _____

12. How much does your family spend on kids' tuition per year (yuan)? _____

13. The level of your kids' living expenses per month (yuan):

1) Below 500 2) 500-1000 3) 1000-2000 4) 2000-3000 5) Above 3000

14. How much does your family spend on the elderly per month (yuan)?

1) Below 500 2) 500-1000 3) 1000-2000 4) Above 2000

15. Cultivated Area (Mu≈0.164 Acre) ____

16. Current Career:

1) Working for enterprise, government and institution 2) Agriculture 3) Start own business (including cab drivers) 4) Retired 5) Part-time jobs 6) Students 7) Others_____ 8) None

17. How much time do you spend on farm work a year? _____months_____days

18. Your monthly income (yuan):

1) Below 500 2) 500-999 3) 1000-1999 4) 2000-2999 5) 3000-3999 6) 4000-4999 7) Above 5000

19. The total income for your family last year (yuan):

1) Below 5,000 2) 5,000-9,999 3) 10,000-25,000 4) 25,000-50,000 5) 50,000-75,000 6) 75,000-100,000 7) Above 100,000 8) Don't Know

20. Planned Usage of your Savings: Use 1 _____ Use 2_____

1) For children 2) Medical treatment 3)For elders 4) For retirement 5)Building 6) Maintain a living 7) For investment 8) For your own marriage 9) Others_____

21. Have you joined the following insurance? _____

1) New Rural Cooperative Medical System 2) Medical Insurance for Urban Workers 3) Commercial Health Insurance 4) None

Part Two: Health Condition and Living Habit

22. Do you smoke?

1) Never 2) Used to smoke occasionally, but now quit 3) Used to smoke every day, but now quit 4) Smoke occasionally now 5) Smoke every day now

23. Do you drink?

1) Never 2) Used to drink occasionally, but now quit 3) Used to drink frequently (3 or more times a week), but now quit 4) Drink occasionally now 5) Drink frequently (3 or more times a week) now

24. You currently sleep _____ hours a day.

25. You typically go to bed at _____ o'clock at night.

26. It usually takes you ____ minutes to fall asleep after you go to the bed.

27. You usually get up _____ o'clock in the morning.

28. Do you often wake up at midnight or wake up too early in the morning? 1) Yes 2) No

29. Do you dream when you are asleep? 1) Very often 2) Occasionally 3) Never

30. How well do you sleep? 1) Good 2) Normal 3) Poor

31. Do you sleep at noon? 1) Yes 2) No

31.1 (If yes) How long do you sleep? _____ Minutes

31.2 In what specific circumstances will you sleep at noon? _____

32. Do you rely on sleeping pills? 1) Yes 2) No

33. Do you snore? 1) Yes 2) No 999) Don't know

34. Your diet structure:

1) Mostly meat 2) Half meat, half vegetables 3) Mostly vegetables 4) Only vegetables

35. What part of eggs do you usually eat?

1) Egg white 2) Yolk 3) The whole egg 4) Don't eat eggs at all

36. How often do you drink milk?

1) Never 2) Rarely (less than 1 to 3 times per month) 3) sometimes (once or twice a week) 4) Often (3-6 times a week) 5) Very often

37. How much staple food do you eat? _____ liang (1 liang = 50 grams)

38. How often do you eat nuts (such as melon seeds, peanuts, chestnuts, walnuts and so on)?

1) Never 2) Rarely (less than 1 to 3 times per month) 3) sometimes (once or twice a week) 4) Often (3-6 times a week) 5) Very often

39. How much vegetable oil does your family eat a month? _____ jin (1 jin = 500 grams)

40. How much lard oil does your family eat a month? _____ jin (1 jin = 500 grams)

41. How much non-starchy vegetables does your family eat? _____ times per _____ (unit: day, week, month, year), with the average amount of _____ grams.

42. So far, what's your peak weight? _____ kilograms

43. So far, your weight reaches its peak at your age of ___.

44. Do you exercise very often? 1) At least 5 times a week 2) 1-3 times a week 3) 1-3 times a month 4) Less than once a month

45. How long does your exercise last?

1) Less than half an hour 2) Half an hour – an hour 3) At least an hour

46. What are your primary methods of exercise?

1) Stroll 2) Brisk walking 3) Jogging 4) Ball games 5) Swimming 6) Dancing outdoors 7) Equipment 8) Others _____ 9) None

47. What type of medicine do you prefer if you don't feel well?

1) Chinese medicine 2) Western medicine 999) Don't know

48. Which hospital would you go if you don't feel well?

1) Community hospitals 2) Hospitals in Pinggu 3) Hospitals in Beijing 999) Don't know

49. Do you trust doctors?

1) Don't trust them at all 2) Don't trust them to some degree 3) Normal 4) Trust them to some degree 5) Trust them very much

50. Have you heard of the harm of diabetes? 1) Yes 2) No

51. In which ways did you hear about diabetes?

1) Radio 2) TV advertising 3) Newspapers and magazines 4) Community activities 5) Relatives and friends 999) Don't know

52. Your knowledge on diabetes:

1) Very familiar 2) Familiar 3)Know a little 4)Know little 5)Know nothing

52.1 Do you think obesity will increase the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.2 Do you think height will affect the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.3 Do you think diet will affect the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.4 Do you think lack of exercises will increase the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.5 Do you think aging will increase the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.6 Do you think blood type will affect the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

52.7 Do you think hypertension will increase the probability of having diabetes?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

53. Do you think diabetes can be inherited?

1) Yes 2) No 999) Don’t know

54. What do you think are the symptoms of diabetes?

1) excessive hunger, excessive thirst, frequent urination and weight loss

2) excessive hunger, excessive thirst, frequent urination and fatigue

3) excessive hunger, excessive thirst, frequent dreams and weight loss 999) Don’t know

55. Diabetes is:

1) Curable 2) Incurable 999) Don’t know

56. Which of the followings can be effective in diabetes treatment?

1) Diet therapy 2) Medical therapy 3) Psychological therapy 4) Sports Therapy 999) Don’t know

57. Do you think diabetes will lead to hypertension, Coronary disease or eye diseases:

1) Yes 2) No 3) Don’t know

58. How difficult is that you keep diet, eat less high sugar food as fruit and sweetmeat?

1) Totally impossible 2) Very difficult 3) A little difficult 4) Basically possible 5) Very possible

59. How difficult is it for you to take medicine every day in order to keep blood glucose level?

1) Impossible 2) Very difficult 3)A little difficult 4) Basically possible 5) Very possible

60. How difficult is it for you to keep exercising and weight every day in order to keep blood glucose level?

1) Impossible 2) Very difficult 3)A little difficult 4) Basically possible 5) Very possible

61. Do you take physical examinations initiatively?

1) Yes 2) No

62. What’s the last time you took a blood sugar test?

1) Never 2) More than 1 year ago 3) Less than 1 year ago 4) Less than 1 month ago 5) Less than half a year ago 6) Less than 3 months ago

62.1 What’s the results of your last blood sugar test?

1) Normal blood sugar 2) Low blood sugar 3) Diabetes 4) Pre-diabetes 5) Forgot it 6) High blood sugar but not diagnosed with diabetes

63. Is there anyone among your grandparents, parents, children and siblings diagnosed with diabetes? 1) Yes 2) No 999) Don’t know

64. Is there any of your relatives diagnosed with diabetes? 1) Yes 2) No 999) Don’t know

65. Is there any of your friends and neighbors diagnosed with diabetes?

1) Yes 2) No 999) Don’t know

66. (If female) Have you given birth to a giant baby? 1) Yes 2) No 999) Forgot it

67. Your subjective possibilities to have diabetes ┬┬┬┬┬┬┬┬┬┬┐_____ (interviewers distribute 10 little balls to interviewees)

68. What do you think is the difference between the diabetes tests that you have access to everyday and the blood tests in large hospitals?

1) The results are more accurate in large hospitals 2) No difference 999) Don't know

69. Do you think cancers can to some degree be controlled?

1) Yes 2) No 999) Don't know

Part Three: Preference

70. Please choose one from the two options in each line.

	Option 1	Option 2	Your choice: 1 or 2?
A	Gain ￥50	Throw the coin. If head shows up, you will be paid ￥200, or you will be paid nothing.	_____
B	Gain ￥80	Throw the coin. If head shows up, you will be paid ￥200, or you will be paid nothing.	_____
C	Gain ￥100	Throw the coin. If head shows up, you will be paid ￥200, or you will be paid nothing.	_____
D	Gain ￥120	Throw the coin. If head shows up, you will be paid ￥200, or you will be paid nothing.	_____
E	Gain ￥150	Throw the coin. If head shows up, you will be paid ￥200, or you will be paid nothing.	_____

71. Please choose one from the two options in each line.

	Option 1	Option 2	Your choice: 1 or 2?
A	Gain ￥10	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50.	_____
B	No Win No Lose	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50	_____
C	Lose ￥10	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50.	_____
D	Lose ￥20	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50.	_____
E	Lose ￥30	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50.	_____

F	Lose ￥40	Throw the coin. If head shows up, you will be paid ￥60, or you will lose ￥50.	_____

72. Please choose one from the two options in each line.

	Option 1	Option 2	Your choice: 1 or 2?
A	￥1000 at present	￥1063 one year later	_____
B	￥1000 at present	￥1188 one year later	_____
C	￥1000 at present	￥1313 one year later	_____
D	￥1000 at present	￥1437 one year later	_____
E	￥1000 at present	￥1563 one year later	_____
F	￥1000 at present	￥1688 one year later73.	_____

73. Please choose one from the two options in each line.

	Option 1	Option 2	Your choice: 1 or 2?
A	￥1000 2 years later	￥1063 3 years later	_____
B	￥1000 2 years later	￥1188 3 years later	_____
C	￥1000 2 years later	￥1313 3 years later	_____
D	￥1000 2 years later	￥1437 3 years later	_____
E	￥1000 2 years later	￥1563 3 years later	_____
F	￥1000 2 years later	￥1688 3 years later	_____

74. Assuming that you got 4 travelling coupons, and each coupon can give you one free travel experience. The coupons are effective from tonight, and will expire in 2 years. Please answer the following questions.

74.1 If you have enough time, in the most ideal circumstances, how would you distribute the coupons?

1) Use 4 coupons in the first year, and 0 in the second year.

2) Use 3 coupons in the first year, and 1 in the second year.

3) Use 2 coupons in the first year, and 2 in the second year.

4) Use 1 coupons in the first year, and 3 in the second year.

5) Use 0 coupons in the first year, and 4 in the second year.

74.2 If you have enough time, and based on your most accurate predict of yourself, how do you think you will distribute the coupons in reality?

1) Use 4 coupons in the first year, and 0 in the second year.

2) Use 3 coupons in the first year, and 1 in the second year.

3) Use 2 coupons in the first year, and 2 in the second year.

4) Use 1 coupons in the first year, and 3 in the second year.

5) Use 0 coupons in the first year, and 4 in the second year.

74.3 If you can get the coupons in two of the following ways, which one would you prefer?

1) 4 coupons at once

2) 4 coupons at once but based on your ideal distribution, some of the coupons can only be used in the second year.

74.4 Assuming that the 4 coupons will become effective after 1 year and period of validity is 2 years, how to distribute the coupons

how would you distribute the coupons?

1) Use 4 coupons in the second year, and 0 in the third year.

2) Use 3 coupons in the second year, and 1 in the third year.

3) Use 2 coupons in the second year, and 2 in the third year.

4) Use 1 coupons in the second year, and 3 in the third year.

5) Use 0 coupons in the second year, and 4 in the third year.

75. You are driving (as a passenger) with an inexperienced and uncertain driver. The weather is very bad, and there is a lot of snow and ice on the road surface.

75.1 (Strategy A) During the process, you will

1) Pay no attention to the driving and road conditions

2) Pay some attention to the driving and road conditions

3) Pay close attention to the driving and road conditions

75.1 (Strategy B) If beautiful music is played in the car, you will

1) Cannot focus on the music, but pay close attention to the driving and road conditions

2) Sometimes listen to the music, and sometimes pay attention to the driving and road conditions

3) Focus on the music and pay no attention to the driving and road conditions

76. For some time, you have complaints about headaches and dizziness. You visit your doctor. The doctor is suspicious about your complaints and sends you to the hospital to undergo an aversive examination.

76.1 (Strategy A) Before taking the examination, you will

1) Not search for any information related to the examination

2) Feel good if there is some related information, and doesn’t matter if there isn’t any

3) Have to find some information related to the examination

76.2 (Strategy B) Before taking the examination, you will

1) Keep thinking about the examination even when doing something else

2) Sometimes think about the examination when doing something else

3) Totally forget the examination when doing something else

77. Late at night, you walk through a deserted neighborhood of a city. Suddenly, a group of dubious looking people approach you from a side-road.

77.1 (Strategy A) You will

1) Keep walking without paying attention to their moves

2) Occasionally take a look at their moves

3) Pay close attention to their moves

77.2 (Strategy B) You will think

1) The reality is as dangerous as it seems

2) The reality may not be as dangerous as it seems

3) The reality is not as dangerous as it seems

78. The Big Five Inventory (BFI)

Here are a number of characteristics that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

Disagree Strongly-1, Disagree a little-2, Neither agree nor disagree-3, Agree a little-4 Agree Strongly-5

78.1 Is depressed, blue

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.2 Is relaxed, handles stress well

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.3 Worries a lot

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.4 Is emotionally stable, not easily upset

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.5 Can be moody

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.6 Remains calm in tense situations

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.7 Gets nervous easily

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.8 I would feel afraid if I had to travel in bad weather conditions.

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

78.9 When it comes to physical danger, I am very fearful.

1) Disagree strongly 2) Disagree a little 3) Neither agree nor disagree 4) Agree a little 5) Agree strongly

Part Four: Take-up Decisions and Physical Examinations

79. The interviewee is asked to draw lots, and the number is recorded without any explanation.

1) Free 2) 10 yuan 3) 30 yuan

80. Are you willing to take the diabetes test? 1) Yes 2) No

[If the interviewee is willing to take the test but don't have enough money, ask this question.]

80.1 Are you willing to take the diabetes test after you fetch the money back home?

1) No, and give up the test. 2) Yes, but change the mind at home 3) Yes, and successfully take the test

81. Why? _____

[If the interviewee is willing to take the test, ask the following questions. If not, move on to Part Five directly.]

82. Are you on an empty stomach? 1) Yes 2) No

[Ask the individuals on an empty stomach to take OGTT and A1C tests. If they don't want to wait for 2 hours, tell them OGTT results will be more accurate. If they still don't want to wait, bring them to take Random Plasma Glucose Test and A1C test. Ask the individuals who are not on an empty stomach to take Random Plasma Glucose Test and A1C test.]

83. Is the interviewee willing to wait 2 hours for OGTT test? 1) Yes 2) No

[Interviewers fill out the following questions by themselves.]

84. How cooperative is the interviewee? 1) Great 2) Good 3) Standard 4) Poor

85. How do you think about this interview 1) Reliable 2) Mostly reliable 3) Not reliable

86. The date of the survey: Year: _____ Month: _____ Date: _____

87. Signature of the interviewer: _____

88. Time the interview end: _____

Part Five: Physical Examinations

89. Your Stature _____cm

90. Your Weight _____kg

91. Your Waistline _____cm

92. Your Hipline _____cm

93. Your Blood Pressure: Systolic Pressure _____mmHg Diastolic Pressure _____mmHg

The survey of Disease experiment is the same, besides that the following questions are added to Part Two and some adjustment in Part Four.

Part Two: Health Condition and Living Habit

69.1 Do you think common cancer can be controlled?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.2 The treatment to which disease do you think will be more expensive?

1) Diabetes 2) Cancer 3) Almost the same 999) Don't know

69.3 Your subjective possibilities of having cancer _____ (interviewers distribute 10 little balls to interviewees)

69.4 Interviewers introduce how diabetes blood test and CEA test works. 1) Yes 2) No

69.5 Do you think diet will affect the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.6 Do you think height will affect the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.7 Do you think smoking will increase the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.8 Do you think lack of exercise will increase the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.9 Do you think a relative having cancer will affect the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.10 Do you think blood type will affect the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.11 Do you think enteritis will increase the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

69.12 Do you think aging will increase the probability of having cancer?

1) Yes (has some effect) 2) No (has almost no effect) 999) Don't know

Part Four: Take-up Decisions and Physical Examinations

79. The interviewee is asked to draw lots, and the number is recorded without any explanation

1) Diabetes treatment 2) Cancer treatment

[If choose 1), skip 80.2; If choose 2), skip 80.1]

80.1 Are you willing to use the existing blood sample to test whether you have diabetes?

1) Yes 2) No

80.2 Are you willing to use the existing blood sample to test whether you have cancer?

1) Yes 2) No

81. Why? _____

82. If you could choose to take one of the tests for free, which one would you prefer?

1) Diabetes test 2) Cancer test 3) Neither or both

Appendix 3: Theory and proofs

3.1 Neoclassical model

There is a binary state $s \in \{0, 1\}$ where $s = 1$ indicates that the individual has the disease (diabetes or cancer) and $s = 0$ otherwise. Individuals have some exogenous $p = E(s)$, which measures the true probability of having the disease. At time 0, individuals choose whether or not to learn the true state through medical testing with cost C. At time 1, individuals choose a binary action $a \in \{0, 1\}$, which can be understood as treatment related to the disease, and experience utility associated with their expectations of time 2 consumption. Ex post individual consumption utility is maximized when action is matched to state. At time 2, the true state is revealed, and individuals receive consumption utility.

When individuals do not learn the true state at time 0, they choose action $\hat{a}(p) = argmax_a E[u(a, s|p)]$.The expected consumption utility at time 2 is $E[u(\hat{a}, s|p)] = pu(\hat{a}, 1) + (1 - p)u(\hat{a}, 0)$. Thus, the utility of not testing is

$$U_{untest} = pu(\hat{a}, 1) + (1 - p)u(\hat{a}, 0)$$

If they learn the true state at time 0, they will adopt ex post optimal action $a = s$. The utility of testing is

$$U_{test} = pu(1,1) + (1 - p)u(0,0)$$

We define the value of testing as the difference between the expected utility if tested and the expected utility if untested, $V_{test} = U_{test} - U_{untest}$. When individuals decide whether to take the test, they compare the utility of testing to the utility of not testing, given their optimal choices.

We define utility function u as follows. Being healthy and taking the state-matched action has a value of 1 (u(0, 0) = 1). Taking the wrong action in the healthy state leads to the utility loss of Φ compared to taking the right action (u(1, 0) =$1 - \Phi$). Being sick and taking the state-matched action has a value of 0 (u(1, 1) = $-v$). $v \geq 0$, and v depends on the type of disease: The more serious the disease, the larger v is. Taking the wrong action in the sick state leads to a utility loss of Ω compared to taking the right action (u(0, 1) = -$\Omega - v$). Therefore Φ measures the cost of taking any action when healthy and Ω measures the cost of not taking any action when sick. We assume that Φ, Ω< 1, which implies that individuals value health more than they value the correct action. When $v = 0$, the model collapses to the model in Oster et al. (2013). We normalize $v = 0$ for diabetes and assume v>0 for cancer, since the utility of cancer patients after taking cancer treatment is likely to be lower than that of diabetes patients after taking diabetes treatment.

We first assume that treatment is not possible without the test. Then, $U_{untest} = pu(0,1) + (1 - p)u(0,0)$. $V_{test} = U_{test} - U_{untest} = p\Omega$. Therefore, the model predicts 100% take-up of test, as illustrated in Figure 7. We then assume that treatment is possible without the test. The optimal solution takes the following form. At time 1, according to Lemma 1 in Oster et al. (2013), $\hat{a}(p) = 0$ if $p \leq \frac{\Phi}{\Phi+\Omega}$ and $\hat{a}(p) = 1$ if $p > \frac{\Phi}{\Phi+\Omega}$. Define $p^* = \frac{\Phi}{\Phi+\Omega} + \frac{\delta\Phi(1+\Omega)}{(\Phi+\Omega)^2}$. At time 0, according to Propositions 1 and 2 in Oster et al. (2013), when individuals remain untested the manipulation of beliefs goes as follows: When $p \leq p^*$, $a = 0$. When $p > p^*$, $a = 1$. The intuition is as follows. When disease

risk is greater than a threshold, the individual becomes more certain about having the disease hence might seek equally-effective alternative treatment even without the test. In this case, when the risk is low and the individual does not seek any treatment, the value of taking the test and receiving the state-matched treatment becomes larger as the risk increases. When the risk is high and the individual already seeks alternative treatment, the value of taking the test becomes smaller as individual the risk increases.

Figure 7 shows that the value of testing is non-monotonic. Thus, there are two potential cutoff points at which the total value of testing is zero. We define p_{low} and p_{high} to be the low and high cutoff points, respectively. Individuals with probability of having the disease lower than p_{low} and higher than p_{high} will avoid the test. We can solve the closed form based on Proposition 3 in Oster et al. (2013). Proposition 1 below summarizes model predictions.

Proposition A1. $p_{low} = \frac{C}{\Omega}$, and $p_{high} = 1 - \frac{C}{\Phi}$ we show that $\frac{\partial p_{low}}{\partial C} > 0$, $\frac{\partial p_{low}}{\partial \Omega} < 0$, $\frac{\partial p_{high}}{\partial C} < 0$, and $\frac{\partial p_{high}}{\partial \Omega} = 0$.

Proof:

$$\frac{\partial p_{low}}{\partial C} = \frac{1}{\Omega} > 0$$

$$\frac{\partial p_{low}}{\partial \Omega} = -\frac{C}{\Omega^2} < 0$$

$$\frac{\partial p_{high}}{\partial C} = -\frac{1}{\Phi} < 0$$

$$\frac{\partial p_{high}}{\partial \Omega} = 0$$

From proposition A1, we can see that the neoclassical model can predict that both the low- and high-risk groups are more likely to avoid the test. This is because their model assumes that high-risk individuals will take the same proper treatment action as in the medical system, even without being formally diagnosed.

The model also predicts that the cutoff point is independent of the utility loss of taking the wrong action when sick ($\frac{\partial p_{high}}{\partial \Omega} = 0$). The intuition is that an individual should always take the test when it is free. Thus, changes in the disease type will not change testing behavior and, all will take the tests.

[Figure A2]

Figure A2 shows the predictions of the neoclassical model for diabetes under the parameter values $\Omega = 0.3$, $\Phi =$

0.3, and $v = 0$. Panel A illustrates the predictions from the price experiment when we vary the cost of testing based on Proposition A1. It shows that both the low- and high-risk groups are more likely to avoid the test as the test cost increases. Panel B illustrates the predictions from the disease experiment when the test is free. For cancer, we choose $v = 0.6$ to represent that individuals with more serious disease have lower health states if not treated. It shows that changes in the disease type will not change testing behavior and all will take the tests.

The above analysis is based on the assumption that equally effective treatment is possible without the test. We can relax the above assumption to imperfect treatment, i.e. treatment is possible without testing but is less effective than the treatment after formal diagnosis for untested individuals. Specifically, if individuals choose to take the test, individuals can choose no action ($a = 0$) or state-match treatment ($a = 1$). If individuals choose not to take the test, individuals can only choose no action ($a = 0$) or imperfect treatment ($a = k$, and $0 < k \leq 1$). When $k = 1$, it becomes the original case. We define the utility function as follows

$$u(k,1) = -(1-k)\Omega,$$

$$u(k,0) = \ (1-k\Phi)$$

Lemma A1: For some $0 < k \leq 1$.

$$a(p) = 0 \ \ if \ \ p \leq \ \frac{\Phi}{\Phi + \Omega}$$

$$and \ \ a(p) = k \ \ if \ \ p > \ \frac{\Phi}{\Phi + \Omega}$$

Proof:

$$pu(0,1) + (1-p)u(0,0) \geq \ \ pu(k,1) + (1-p)u(k,0)$$

$$-p\Omega + (1-p) \geq \ \ -p(1-k)\Omega + (1-p)(1-k\Phi)$$

$$\frac{\Phi}{\Phi + \Omega} \geq \ \ p$$

We can derive the following utility function.

$$U = \begin{cases} (1-p) - p\Omega & if \ \ p \leq \ \frac{\Phi}{\Phi + \Omega} \\ (1-p)(1-k\Phi) - p(1-k)\Omega & if \ \ p > \ \frac{\Phi}{\Phi + \Omega} \end{cases}$$

And then show the following propositions.

Proposition A2: $p_{low} = \frac{C}{k\Omega}$, and $p_{high} = \ \frac{k\ \phi - C}{k\ \phi - (1-k)\Omega}$, we show that $\frac{\partial p_{low}}{\partial C} > 0$, and $\frac{\partial p_{high}}{\partial C} < 0$

Proof: $\frac{\partial p_{low}}{\partial C} = \frac{1}{k\Omega} > 0$, $\frac{\partial p_{high}}{\partial C} = -\frac{1}{k\ \phi - (1-k)\Omega} \ < 0$

Proposition A2 is similar to Proposition A1. It shows that we can relax the assumption that equally-effecetive

treatment is possible without the test to imperfect treatment.

3.2 Optimal expectation model

The key assumption in the model is that individuals experience anticipation utility over future health status. Individuals form beliefs about their probability of having the disease, π, and π can be different from the true probability p. Let $u(a, s)$ be the consumption utility given action a and health state s. Let δ be the weight on anticipation utility. The following equation gives the utility function at time 0, which is a weighted average of anticipation utility based on π and consumption utility based on p.

$$U(\pi|p) = \delta E[u(\hat{a}, s|\pi)] + E[u(\hat{a}, s|p)]$$

The optimal choices are derived in a backward-induction manner. At time 1, individuals decide on the optimal action given belief π to maximize anticipatory utility. At time 0, individuals maximize $U(\pi|p)$ by choosing whether to take the test, and if not, what is the optimal belief π. If individuals take the test, their beliefs become rational, so π=p. If they remain untested, they also choose the optimal π to maximize total utility $U(\pi|p)$.

When individuals do not learn the true state at time 0, they choose action $\hat{a}(\pi) = argmax_a E[u(a, s|\pi)]$. In this case, the anticipation utility at time 1 is $\delta E[u(\hat{a}, s|\pi)] = \delta(\pi u(\hat{a}, 1) + (1-\pi)u(\hat{a}, 0))$, and the expected consumption utility at time 2 is $E[u(\hat{a}, s|p)] = pu(\hat{a}, 1) + (1-p)u(\hat{a}, 0)$. Thus, the utility of not testing is

$$U_{untest} = \delta\big(\pi u(\hat{a}, 1) + (1-\pi)u(\hat{a}, 0)\big) + pu(\hat{a}, 1) + (1-p)u(\hat{a}, 0)$$

If they learn the true state at time 0, they will adopt ex post optimal action $a = s$. The utility of testing is

$$U_{test} = (1+\delta)[pu(1,1) + (1-p)u(0,0)]$$

When individuals decide whether to take the test, they compare the utility of testing to the utility of not testing, given their optimal choices.

We first consider the price experiment with diabetes tests. The optimal solution takes the following form. At time 1, according to Lemma 1 in Oster et al. (2013), $\hat{a}(\pi) = 0$ if $\pi \leq \frac{\Phi}{\Phi+\Omega}$ and $\hat{a}(\pi) = 1$ if $\pi > \frac{\Phi}{\Phi+\Omega}$. Define $p^* = \frac{\Phi}{\Phi+\Omega} + \frac{\delta\Phi(1+\Omega)}{(\Phi+\Omega)^2}$. At time 0, according to Propositions 1 and 2 in Oster et al. (2013), when individuals remain untested the manipulation of beliefs goes as follows: When $p \leq p^*$, $\pi = 0$ and $a = 0$. When $p > p^*$, $\pi = \frac{\Phi}{\Phi+\Omega}$ and $a = 1$. The intuition is as follows. Since action is binary, there is a range in which changing π does not change the optimal actions, and hence the consumption utility. To maximize anticipation utility, individuals will choose the lowest π in that range, leading to the corner solution of π.

Figure 8 shows that the value of testing is non-monotonic. Thus, there are two potential cutoff points at which the total value of testing is zero. We define p_{low} and p_{high} to be the low and high cutoff points, respectively.

Individuals with probability of having the disease lower than p_{low} and higher than p_{high} will avoid the test. We can solve the closed form based on Proposition 3 in Oster et al. (2013). Proposition 1 below summarizes model predictions.

Proposition A3 (Price experiment). When $0 < \delta + \delta v < \Omega$, $p_{low} = \frac{C}{\Omega - \delta - \delta v}$, and $p_{high} = \frac{\delta\Phi(1+\Omega+v)}{(\Phi+\Omega)(\delta+\delta v+\Phi)} + \frac{\Phi - C}{\delta+\delta v+\Phi}$, we show that $\frac{\partial p_{low}}{\partial C} > 0$, and $\frac{\partial p_{high}}{\partial C} < 0$.

Proof: According to Propositions 1 in Oster et al. (2013), when choosing not to take the test, as $> p^*$, $\pi = \frac{\Phi}{\Phi+\Omega}$. For the cancer treatment, when $u(1,1) = -v$, and $v>0$, $\pi = \frac{\Phi}{\Phi+\Omega_c}$ and $p^* = \frac{\Phi}{\Phi+\Omega_c} + \frac{\delta\Phi(1+\Omega_c)}{(\Phi+\Omega_c)^2}$. The cutoff point is the following

$$p_{low} = \frac{C}{\Omega - \delta - \delta v}$$

$$p_{high} = \frac{\delta\Phi(1+\Omega+v)}{(\Phi+\Omega)(\delta+\delta v+\Phi)} + \frac{\Phi - C}{\delta+\delta v+\Phi}$$

$$\frac{\partial p_{low}}{\partial C} = \frac{1}{\Omega - \delta - \delta v} > 0$$

$$\frac{\partial p_{high}}{\partial C} = -\frac{1}{\delta+\delta v+\Phi} < 0$$

Proposition A3 explores how the cutoff level for low- and high-risk individuals to avoid taking the test varies with test price (C). We focus on p_{high}, since this is more relevant for our purpose. Proposition A3 suggests that the higher the test price, the lower the high cutoff point, and the more likely we will observe information avoidance in the high-risk group, given that observations at the extreme right tail may be scarce empirically. The intuition is that given the non-monotonic relationship in Figure 8, an increase in the test price will reduce the value of testing, and thus marginal individuals around high cutoffs will avoid the information.

In the price experiment, we randomize the price of diabetes testing, which is the cost of test *C* in the model. In the disease experiment, since cancer is more serious than diabetes, it is reasonable to assume that individuals would incur more loss from cancer if they do not treat the disease when they have it.

Figure A3, Panel A illustrates the predictions from the price experiment when we vary the cost of testing based on Proposition A3. For diabetes, we choose $\delta = 0.2$, $\Omega = 0.3$, $\Phi = 0.3$, and $v = 0$. The horizontal axis is *p*, the vertical axis is the total value of testing, and *C* is the cost of testing. When *C*=0, the total value of testing is an inverse V-shape over *p*, which is the same as Figure 8. In this case, those with low *p* are likely to take the test and those with high *p* will not take the test. When *C*>0, the total value of testing moves downward. In this case, those with very low *p* and very high *p* are predicted to not take the test. The model, therefore, makes the following two predictions for the

price experiment.

[Figure A3]

Prediction 1: When the test price is positive, the relationship between beliefs about disease probability and test take-up is non-monotonic: The take-up rate should be lower for low- and high-risk groups.

The take-up for the high-risk group is likely to be low due to the benefit of holding biased beliefs. For the low-risk group the take-up is low, since the benefit of testing to avoid utility loss is also low. This prediction provides a reasonable explanation for the observed non-monotonic pattern in T30 (finding 1). However, the patterns in T0 and T10 are not fully consistent with this prediction. One possible reason is the following: Proposition 1 suggests that the predicted cutoff point for the high-risk group to avoid the test decreases with test cost. For very low test cost, such as our T0 and T10, the cutoff point for the high-risk group to avoid the test is very high. In reality. there may be few observations with subjective risk beyond this high cutoff point. With the higher price in T30, however, the cutoff point is not that high, so we have observations beyond that point to demonstrate the non-monotonic pattern empirically. The pattern for T30 is also more informative, because it is the most comparable to the market price, and hence this non-monotonic pattern is more relevant.

Prediction 2: In the price experiment, increasing the test price will reduce take-up for both the low-risk and the high-risk groups. Thus, conditioning on taking the tests, increasing the test price will reduce the dispersion of test outcomes.

Increasing the test price will reduce the total value of testing. Following prediction 1, when the take-up for tests is an inverse V-shape over p, marginal individuals who take the test in high p would not choose to take the test. Marginal individuals who take the test in low p would also not choose to take the test. Therefore, increasing the test price is likely to keep the average test outcomes unchanged, and it will also reduce the dispersion of test outcomes. This is also consistent with our finding 2.

We then derive the prediction for the disease experiment. Our model differs in two ways from that of Oster et al. (2013). First, we assume $v>0$, since the utility of cancer patients after taking cancer treatment is likely to be lower than that of diabetes patients after taking diabetes treatment. Second, the utility loss of the wrong action relative to the right action given cancer, represented by Ω_c, is higher than when individuals have diabetes, represented by Ω. Therefore we define the utility level of taking the wrong action given cancer as $u(0,1) = -v - \Omega_c$, assuming that $\Omega_c > \Omega$.

Proposition A4 (Disease experiment). When $0 < \delta + \delta v < \Omega$, $p_{low} = \frac{C}{\Omega - \delta - \delta v}$, and $p_{high} = \frac{\delta\Phi(1+\Omega+\mathrm{v})}{(\Phi+\Omega)(\delta+\delta v+\Phi)} + \frac{\Phi - C}{\delta+\delta v+\Phi}$, We show that $\frac{\partial p_{low}}{\partial \Omega} < 0$, $\frac{\partial p_{low}}{\partial \mathrm{v}} > 0$, $\frac{\partial p_{high}}{\partial \Omega} < 0$, and $\frac{\partial p_{high}}{\partial \mathrm{v}} < 0$ when C is small.

Proof: Similar to Proposition 1, the cutoff point is the following

$$p_{low} = \frac{C}{\Omega - \delta - \delta v}$$

$$p_{high} = \frac{\delta\Phi(1+\Omega+\mathrm{v})}{(\Phi+\Omega)(\delta+\delta v+\Phi)} + \frac{\Phi - C}{\delta+\delta v+\Phi}$$

$$\frac{\partial p_{low}}{\partial \Omega} = -\frac{C}{(\Omega - \delta - \delta v)^2} < 0$$

$$\frac{\partial p_{low}}{\partial \mathrm{v}} = \frac{\delta C}{(\Omega - \delta - \delta v)^2} > 0$$

$$\frac{\partial p_{high}}{\partial \Omega} = \frac{\delta\Phi(\Phi+\Omega)(\delta+\delta v+\Phi) - \delta\Phi(1+\Omega+\mathrm{v})(\delta+\delta v+\Phi)}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2}$$

$$= \frac{\delta\Phi(\delta+\delta v+\Phi)[(\Phi+\Omega) - (1+\Omega+\mathrm{v})]}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2}$$

$$= \frac{\delta\Phi(\delta+\delta v+\Phi)(\Phi - 1 - \mathrm{v})}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2} < 0$$

$$\frac{\partial p_{high}}{\partial \mathrm{v}} = \frac{\delta\Phi(\Phi+\Omega)(\delta+\delta v+\Phi) - \delta\Phi(1+\Omega+\mathrm{v})[\delta(\Phi+\Omega)]}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2} - \frac{\delta(\Phi - C)}{(\delta+\delta v+\Phi)^2}$$

$$= \frac{\delta\Phi(\Phi+\Omega)[(\delta+\delta v+\Phi) - \delta(1+\Omega+\mathrm{v})]}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2} - \frac{\delta(\Phi - C)}{(\delta+\delta v+\Phi)^2}$$

$$= \frac{\delta\Phi(\Phi+\Omega)(\Phi - \delta\Omega)}{(\Phi+\Omega)^2(\delta+\delta v+\Phi)^2} - \frac{\delta(\Phi - C)}{(\delta+\delta v+\Phi)^2} = \frac{\delta[\Phi(\Phi - \delta\Omega) - (\Phi+\Omega)(\Phi - C)]}{(\Phi+\Omega)(\delta+\delta v+\Phi)^2}$$

$$= \frac{\delta[C\Omega - \Phi(\Omega - \mathrm{C}) - \Phi\delta\Omega]}{(\Phi+\Omega)(\delta+\delta v+\Phi)^2}$$

Thus, $\frac{\partial p_{high}}{\partial \mathrm{v}} = \frac{\delta[C\Omega - \Phi(\Omega - \mathrm{C}) - \Phi\delta\Omega]}{(\Phi+\Omega)(\delta+\delta v+\Phi)^2} < 0 \ \ if \ \ C < \frac{\Phi\Omega(1+\delta)}{\Phi+\Omega}$

Proposition A4 explores how the cutoff level for low- and high-risk individuals to avoid taking the test varies with the utility loss of taking the wrong action when being sick (Ω) and how serious the disease is (v). Figure A3, Panel B illustrates the prediction from the disease experiment when the test is free. For cancer, we choose $v = 0.6$ to represent that individuals with more serious disease have lower health states if not treated. Given the two changes in v and Ω, we see that the non-monotonic relationship between *p* and the value of screening is stronger in the cancer

group, i.e., the cutoff level of p beyond which high-risk individuals will choose not to take the test is lower in the case of cancer than diabetes. Proposition A4 shows that both changes in parameters contribute to the stronger non-monotonic relationship. Each change alone also has the same prediction. The intuition is as follows. First, since the health state with cancer is worse than with diabetes, even after the proper treatment ($v > 0$), the benefit of taking action when sick is lower for cancer. This reduces the benefit of taking the test. Second, since cancer has higher Ω, individuals will take the proper action ($a = 1$) without taking the test even when π is lower, which allows these individuals to hold more optimistic beliefs (lower π) while still avoiding the cost of not taking action when sick. This increases the benefit of not taking the test, and thus marginal individuals around high cutoffs will avoid the information. We therefore have the following prediction:

Prediction 3: In the disease experiment, since the cancer test has larger v and Ω than the diabetes test, the non-monotonic relationship between beliefs about disease risk and take-up for tests is stronger in the cancer treatment.

Note that Proposition A4 suggests that the high cutoff point is decreasing in Ω. In practice there might not be many people with extremely high beliefs about disease risk, so we might not observe the non-monotonic relationship due to lack of observations at the right tail. The cancer treatment strengthens the non-monotonic relationship, because the cutoff point for the high-risk group to avoid the test is lower and therefore we are likely to observe the right-tail pattern of not taking the test. This prediction can explain finding 3.

Regarding finding 4, the explanation depends on how we interpret "the disease is less controllable" in terms of the change in parameters in the model. There are three possibilities. Either "less controllable" means higher v—i.e., the health status is worse even after the treatment—or it means lower Ω —i.e., the utility cost from the wrong action (no treatment) when sick is low—or both. Higher v alone predicts that "less controllable" belief implies more salient information avoidance in the high-risk group, but lower Ω predicts the opposite. If both parameters change, our simulation suggests that as long as the change in v is not too drastic, the v effect dominates and the model prediction is consistent with our finding 4. Therefore, the model can still predict a reasonable explanation for finding 4. We do not observe this pattern in diabetes treatment. This is likely due to the fact that diabetes is in general less severe than cancer so that even less controllable diabetes is still less severe.

- **The model with anticipatory utility and imperfect treatment for untested individuals**

If individuals choose to take the test, individuals can choose no action ($a = 0$) or state-match treatment ($a = 1$). If individuals choose not to take the test, individuals can only choose no action ($a = 0$) or imperfect treatment ($a = k$, and $0 < k \leq 1$). When $k = 1$, it becomes the original case.

$$u(k,1) = -(1-k)\Omega - v,$$

$$u(k,0) = \ (1-k\Phi)$$

Lemma A2: For some $0 < k \leq 1$.

$$a(\pi) = 0 \ \ if \ \ \pi \leq \ \frac{\Phi}{\Phi+\Omega}$$

$$and \ \ a(\pi) = k \ \ if \ \ \pi > \ \frac{\Phi}{\Phi+\Omega}$$

Proof:

$$\pi u(0,1) + (1-\pi)u(0,0) \geq \ \pi u(k,1) + (1-\pi)u(k,0)$$

$$-\pi(\Omega+v) + (1-\pi) \geq \ -\pi((1-k)\Omega+v) + (1-\pi)(1-k\Phi)$$

$$\frac{\Phi}{\Phi+\Omega} \geq \ \pi$$

Proposition A5: Individuals will always choose belief such that $\pi \leq p$.

$$U = \begin{cases} \delta(1-\pi) + (1-p) - (\delta\pi + p)(\Omega+v) & if \ \ \pi \leq \ \frac{\Phi}{\Phi+\Omega} \\ \big(\delta(1-\pi) + (1-p)\big)(1-k\Phi) - (\delta\pi+p)((1-k)\Omega+v) & if \ \ \pi > \ \frac{\Phi}{\Phi+\Omega} \end{cases}$$

Proof: We have assumed that $\Phi, \Omega < 1$, so the agent will only ever choose either $\pi = 0$ or $\pi = \frac{\Phi}{\Phi+\Omega}$. As long as the cutoff point at which individuals switch to belief $\pi = \ \frac{\Phi}{\Phi+\Omega}$ is above $p = \frac{\Phi}{\Phi+\Omega}$, we then have results that $\pi \ll p$. Individuals will choose $\pi = 0$ if the following inequality holds:

$$\delta + \ (1-p) - p(\Omega+v) \geq \ \left(\delta\left(\frac{\Omega}{\Phi+\Omega}\right) + (1-p)\right)(1-k\Phi) - \left(\delta\left(\frac{\Phi}{\Phi+\Omega}\right) + p\right)((1-k)\Omega+v)$$

$$p \leq \frac{\delta\Phi(\Omega+v+1)}{k(\Phi+\Omega)^2} + \frac{\Phi}{\Phi+\Omega}$$

Since $p^* = \ \frac{\delta\Phi(\Omega+v+1)}{k(\Phi+\Omega)^2} + \frac{\Phi}{\Phi+\Omega} > \frac{\Phi}{\Phi+\Omega}$, it follows that $\pi \ll p$.

Proposition A6 Action $a = 0$ will be chosen for values of $p \leq p^*$, and action $a = k$ will be chosen for values of $p > p^*$.

Proof: By Proposition 1, if $p \leq p^*$, then the agent will choose $\pi = 0$, then by Lemma 1, the agent would choose $a = 0$. If $p > p^*$, then the agent will choose $\pi = \frac{\Phi}{\Phi+\Omega}$, then by Lemma 1, the agent would choose $a = k$.

We calculate the potential cutoff points at which the total value of testing is zero, where $U_{test} - \ U_{untest} = 0$.

Note that $U_{test} - \ U_{untest} > 0,$ means individual will choose to take the test.

Where the utility of not testing is

$$U_{untest} = \delta\big(\pi u(\hat{a},1) + (1-\pi)u(\hat{a},0)\big) + pu(\hat{a},1) + (1-p)u(\hat{a},0)$$

and the utility of testing is

$$U_{test} = (1+\delta)[pu(k,1) + (1-p)u(0,0)] - C$$

If $p \le p^*, \pi = 0, a = 0,$

$$(1+\delta)[-p((1-k)\Omega + v) + (1-p)] - C - \delta - (1-p) + p(\Omega + v) = 0$$

$$p_{low} = \frac{C}{(k-(1-k)\delta)\Omega - \delta v - \delta}$$

If $p > p^*, \pi = \frac{\Phi}{\Phi+\Omega}, a = k,$

$$(1+\delta)[-p((1-k)\Omega + v) + (1-p)] - C - \left(\delta\left(\frac{\Omega}{\Phi+\Omega}\right) + (1-p)\right)(1-k\Phi) + \left(\delta\left(\frac{\Phi}{\Phi+\Omega}\right) + p\right)((1-k)\Omega + v) = 0$$

$$p_{high} = \frac{\delta\Phi(1+\Omega+v)}{(\Phi+\Omega)\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)} + \frac{k\Phi - C}{\delta+\big((1-k)\Omega+v\big)\delta+k\Phi}$$

Proposition A7 (Price experiment): When $(k-(1-k)\delta)\Omega > \delta v + \delta > 0$, $p_{low} = \frac{C}{(k-(1-k)\delta)\Omega-\delta v-\delta}$, and $p_{high} = \frac{\delta\Phi(1+\Omega+v)}{(\Phi+\Omega)\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)} + \frac{k\Phi-C}{\delta+\big((1-k)\Omega+v\big)\delta+k\Phi}$, we show that $\frac{\partial p_{low}}{\partial C} > 0$, and $\frac{\partial p_{high}}{\partial C} < 0$

Proof: $\frac{\partial p_{low}}{\partial C} = \frac{1}{(k-(1-k)\delta)\Omega-\delta v-\delta} > 0$, $\frac{\partial p_{high}}{\partial C} = -\frac{-1}{\delta+\big((1-k)\Omega+v\big)\delta+k\Phi} > 0$

Proposition A8 (Disease experiment): When $(k-(1-k)\delta)\Omega > \delta v + \delta > 0$, $p_{low} = \frac{C}{(k-(1-k)\delta)\Omega-\delta v-\delta}$, and $p_{high} = \frac{\delta\Phi(1+\Omega+v)}{(\Phi+\Omega)\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)} + \frac{k\Phi-C}{\delta+\big((1-k)\Omega+v\big)\delta+k\Phi}$, $\frac{\partial p_{low}}{\partial \Omega} < 0$, $\frac{\partial p_{low}}{\partial \mathrm{v}} > 0$, $\frac{\partial p_{high}}{\partial \Omega} < 0$, and $\frac{\partial p_{high}}{\partial \mathrm{v}} < 0$ if C is small.

Proof: $\frac{\partial p_{low}}{\partial \Omega} = -\frac{C(k-(1-k)\delta)}{((k-(1-k)\delta)\Omega-\delta v-\delta)^2} < 0$

$\frac{\partial p_{low}}{\partial \mathrm{v}} = \frac{\delta C}{((k-(1-k)\delta)\Omega-\delta v-\delta)^2} > 0$

$\frac{\partial p_{high}}{\partial \Omega} = \frac{\delta\Phi\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)(\Phi-1-v)-\delta(1-k)(\Phi+\Omega)[\delta\Phi(1+\Omega+v)+(k\Phi-C)(\Phi+\Omega)]}{(\Phi+\Omega)^2\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)^2} < 0$

$\frac{\partial p_{high}}{\partial \mathrm{v}} = \frac{\delta\big(C\Omega-\big((1+\delta)k\Omega-C\big)\Phi\big)}{(\Phi+\Omega)\big(\delta+\big((1-k)\Omega+v\big)\delta+k\Phi\big)^2} < 0$ iff $C < \frac{k\Omega\Phi(1+\delta)}{\Phi+\Omega}$ $\forall 0 < k \le 1$

- **Discussion about objective probability (p) and subjective probability (π)**

In our empirical analysis, we measure subjective probabilities of have diabetes or cancer and analyze the relationship between testing and subjective probabilities. Strictly speaking the subject probabilities correspond to π in the model. According to Appendix Propositions 1 in Oster et al. (2013), when action space is continuous, $\pi =$

$\frac{2p\Omega-\delta(\Omega+1)}{2\Omega(1-\delta)}$. Thus, π is an increasing function of true probability *p*. Thus, high π implies high *p*, our predictions 1-3 remain the same with *p* replaced by π.

3.3 The Attention Model

The basic decision problem is illustrated as a simple three-period timeline. At *t* =0 the person has a prior μ_0, and chooses whether to receive the information. At *t*=1, the person receives information *s*, updates beliefs to μ_1, experiences *t*=1 consumption utility (u_p, which is taken to be exogenous "present utility"), experiences anticipatory utility, and possibly takes an action (x). At *t*=2, state (ω) is revealed and the person experiences *t*=2 consumption utility. The total utility function has three components: attention-based anticipation utility, belief-based anticipation utility, and expected consumption utility:

$$E_{\mu_1}\left[\mathrm{D(t,0)}\Sigma_{i=1}^{I} a(\mu_0,\mu_1)\cdot\left[E_{\mu_1}\left[u\left(x(\omega)\right)\right]-u_p\right]+\mathrm{D(t,1)}\upsilon(\mu_1)+\mathrm{D(t,2)}E_{\mu_1}\left[u\left(x(\omega)\right)\right]\right]$$

Attention-based anticipation utility is defined as $\Sigma_{i=1}^{I} a(\mu_0,\mu_1)\cdot[E_{\mu_1}[\mathrm{u}\left(x(\omega)\right)]-u_p]$. a represents the attention function. The key assumption regarding the attention function is that $a>0$ if $E_{\mu_1}[\mathrm{u}\left(x(\omega)\right)]\neq E_{\mu_0}[\mathrm{u}\left(x(\omega)\right)]$; otherwise, $a=0$. The intuition is that if a person receives information that is utility relevant to future experience, then a boost in anticipatory utility is felt immediately and is proportional to the difference between expected consumption utility and the exogenous present utility, $a[E_{\mu_1}[\mathrm{u}\left(x(\omega)\right)]-u_p]$. $\upsilon(\mu_1)$ is the belief-based anticipation utility based on Caplin and Leahy (2001), which is a von Neumann-Morgenstern utility function directly over the level of posterior beliefs. This belief-based anticipatory utility is in spirit the same as the key component in the curvature model we will discuss in Section 4.4; therefore, we do not focus on this aspect in this section. $E_{\mu_1}[u\left(x(\omega)\right)]$ is the standard consumption utility.

The asterisk line in Figure 9 shows that the negative utility shock reduces the value of testing, and the decrease in value of testing is larger as risk increases, making high-risk individuals avoid the test even when the test has zero cost. The intuition is as follows. The attention-based anticipatory utility is declining in *p*. This is because high-risk patients have a lower expected period 2 outcome $E_{\mu_1}[u\left(x(\omega)\right)]$, and with the assumption that period 1 consumption utility u_p is exogenous, the total value of $E_{\mu_1}[u\left(x(\omega)\right)]-u_p$ is more negative—i.e., This is because high-risk individuals experience more negative shock from learning their test outcomes: the expected utility is lower in the future while current utility is independent of the disease risk.

[Figure A4]

Figure A4, Panel A illustrates the predictions from the price experiment when we vary the cost of testing. For

diabetes, we choose $a(\mu_0, \mu_1)$=0.1, $u_p = 1, \Omega = 0.3, \Phi = 0.3$. The horizontal axis is *p*, the vertical axis is the total value of testing, and *C* is the cost of testing. When *C*=0, the total value of testing is an inverse V-shape over *p*, which is the same as Figure 8. In this case, those with low *p* are likely to take the test and those with high *p* will not take the test. When *C*>0, the total value of testing moves downward. In this case, those with very low *p* and very high *p* are predicted to not take the test. Panel B illustrates the prediction from the disease experiment when the test is free. For cancer, we choose $a(\mu_0, \mu_1)$=0.25 to represent that individuals give more attention to more serious disease. We see that the non-monotonic relationship between *p* and the value of screening is stronger in the cancer group, i.e., the cutoff level of *p* beyond which high-risk individuals will choose not to take the test is lower in the case of cancer than diabetes.

3.4 The Curvature Model

There are two periods, 1 and 2, and individual utility is only derived from health outcomes in period 2. In period 1, the individual forms belief *s* about her health status, which can take any value in the interval (*A*, *B*), with probability distribution function *f*(*s*) and cumulative distribution function *F*(*s*). In period 1, she decides whether to learn about *s* with the cost of tests *c* or remain uninformed. We assume that she learns *s* precisely with the test. After learning *s*, she chooses *t*, which represents treatment from the doctor or a health-relevant lifestyle choice she makes. In period 2, the individual learns her state of health, which depends on *s* and *t* in period 1.

Since we study the screening behavior in period 1, we focus on the anticipatory utility in period 1. The main component of the information avoidance model is that the utility the patient maximizes in the first-period utility derives from her expected belief about her health status in the second period.

$$u(E[s - l(s,t)|\text{patient's information}]) \quad (5)$$

$l(s, t)$ is a loss function that captures the loss of utility if *t* is not the optimal action given *s*. $l(s, t) \geq 0$ for all *s*, *t*, and $\min_{t \in T} l(s,t) = 0$. The patient's utility depends on her expected health in period 2 conditional on her information. Note that anticipatory utility is not defined as expected utility over health status; instead, it is defined as the utility over expected health status, which demonstrates the idea that beliefs generate utility.

Given that the expected belief about health status enters the utility function, the shape of *u* determines the patient's preferences for information. If *u* is concave, she is called "information averse." The intuition is that concave *u* implies diminishing marginal utility over belief about health status. If the individual decides to learn about her health status in period 1, there would be gains and losses compared to the expected value. The loss from learning bad news (moving beliefs to a low health status) is larger than the gain from learning good news (moving beliefs to a

high health status). Thus, the utility from expected health status (avoiding the news) is higher than the expected utility over health status (learning the news). The intuition is similar to risk aversion, where utility is defined over wealth. The key difference is that the utility is defined over expected belief in this framework. Similarly, if u is linear, she is "information neutral;" if u is convex, she is "information loving."

Let t^* represent the optimal treatment given the belief. If she chooses to take the test and update s to be the degenerate health status, she chooses $t = s$ in each state of the world, which gives her an expected utility of $E[u(s)]$. When the patient does not learn the signal on s, she chooses $t^* = \text{argmin}_t E[l(s,t)]$ to minimize the expected loss in health. In this case, her expected utility is the utility on expected health, $u(E[s - l(s,t^*)])$.

The patient prefers to learn her state of health if the utility of testing is greater than the utility of not testing.

$$E[u(s)] > u(E[s - l(s,t^*)]) \quad (6)$$

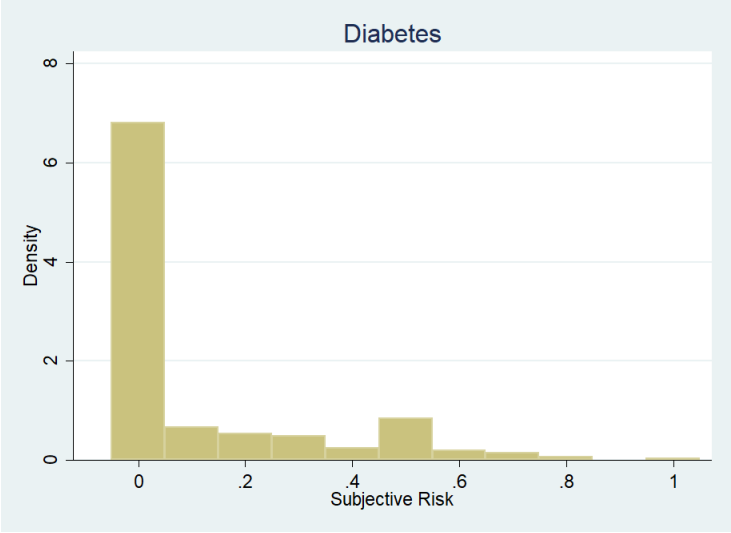
[Figure A5]

Figure A5, Panel A illustrates the predictions from the price experiment when we vary the cost of testing. In our simulation, we assume that $u(s) = s^{\alpha}$. There are two health states: low(s_l) and high (s_h). For diabetes, we choose $\alpha = 0.8$, $s_l = 1$, $s_h = 2$. When individuals do not take the test, $l(s,t^*) = \begin{cases} 0.05 & if \ s = s_l \\ 0 & if \ s = s_h \end{cases}$. The horizontal axis is p, the vertical axis is the total value of testing, and C is the cost of testing. When C=0, those with low p and high p will not take the test. When C>0, the total value of testing moves downward. In this case, more individuals with very low p and very high p will not take the test. Panel B illustrates the prediction from the disease experiment when the test is free. For cancer, we choose $s_l = 0.9$ to represent that individuals with more serious disease have lower health states in low state. We see that the non-monotonic relationship between p and the value of screening is stronger in the cancer group, i.e., the cutoff level of p beyond which high-risk individuals will choose not to take the test is lower in the case of cancer than diabetes.

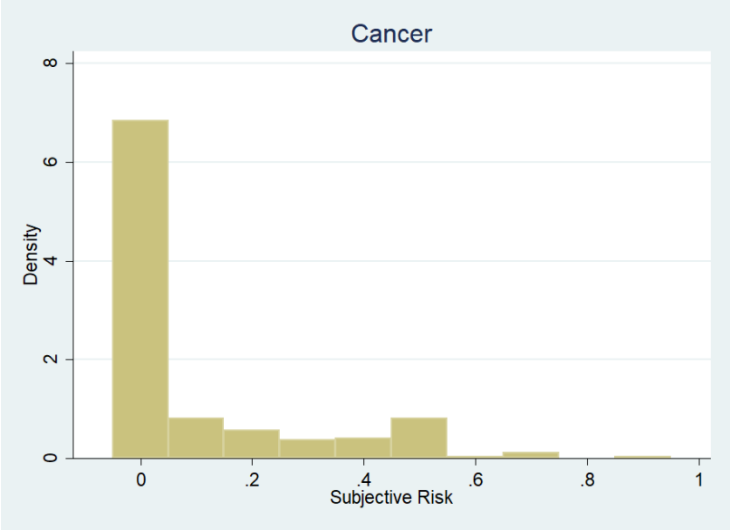
Appendix 4. Other Tables and Figures

Figure A1. Distribution of Subjective Risk

Panel A. Diabetes



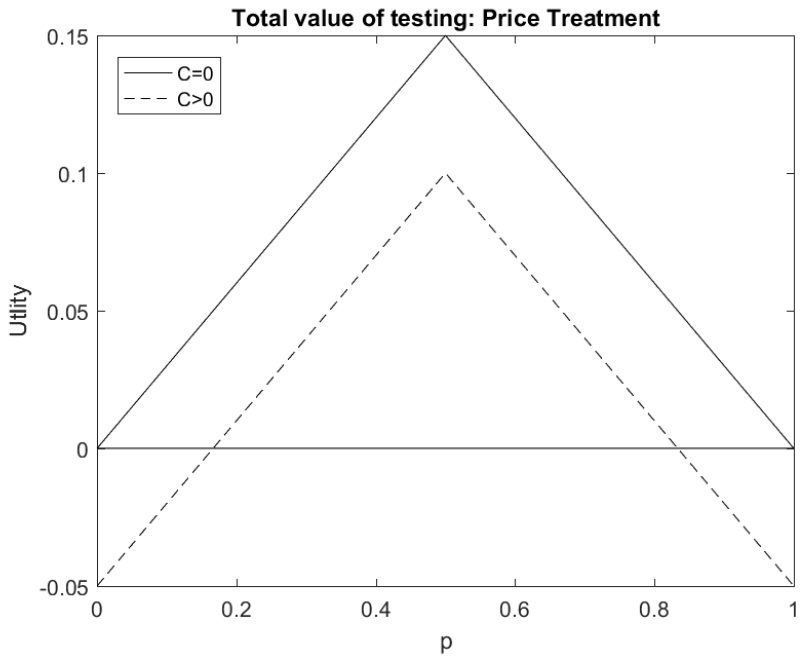
Panel B. Cancer



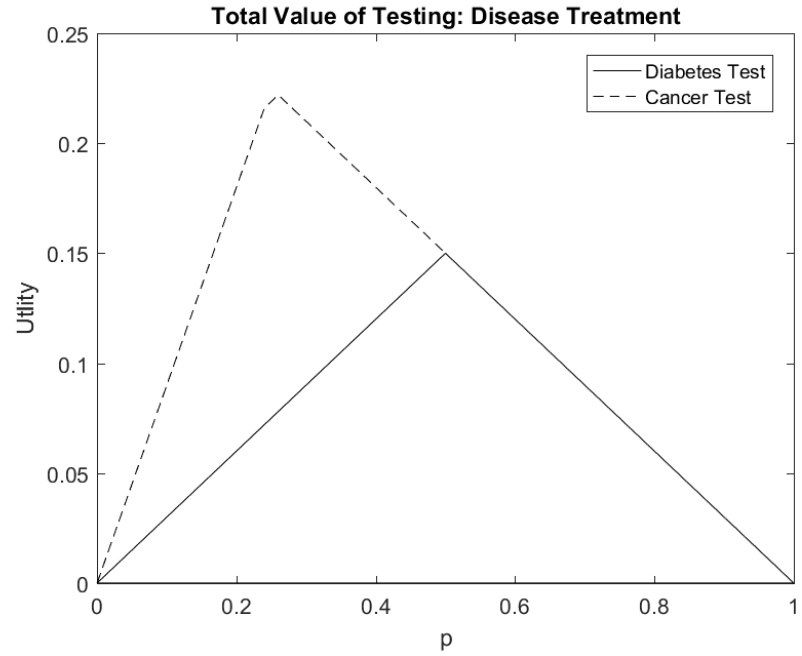
Note: Panel A shows the distribution of subjective risk for diabetes. Panel B shows the distribution of subjective risk for cancer.

Figure A2. Predictions for Price experiment and Disease experiment: Neoclassical Model

Panel A. Price experiment



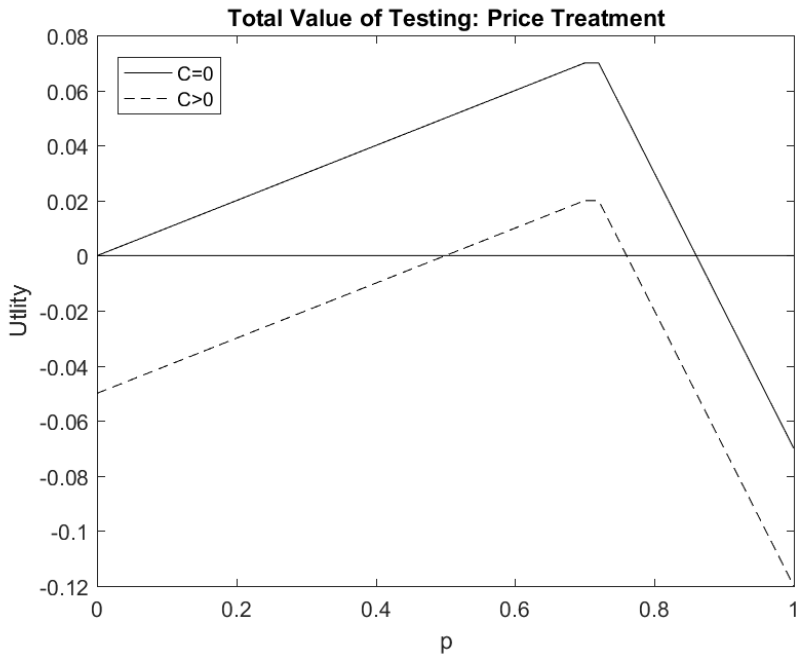
Panel B. Disease experiment



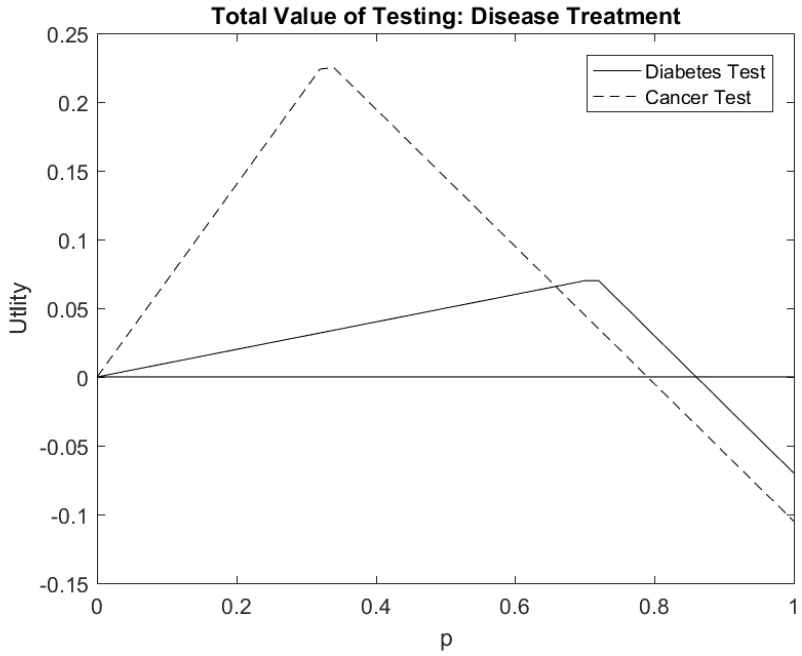
Note: These figures show the predictions of the model with no anticipatory utility. Panel A illustrates the predictions from the price experiment when we vary the cost of testing. For diabetes, we choose $\Omega = 0.3$, $\Phi = 0.3$, and $v = 0$. The horizontal axis is p. The vertical axis is the total value of testing. C is the cost of testing. Panel B illustrates predictions from the disease experiment when the test is free. For cancer, we choose $v = 0.6$ to represent that individuals with more serious disease have lower health states if not treated.

Figure A3. Predictions of the Optimal Expectations Model for the Price and Disease experiment

Panel A. Price experiment



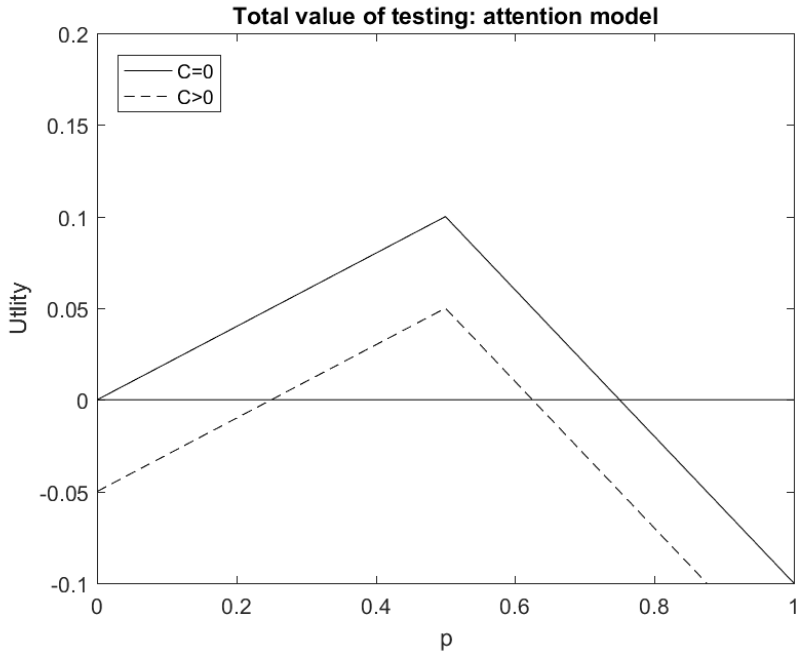
Panel B. Disease experiment



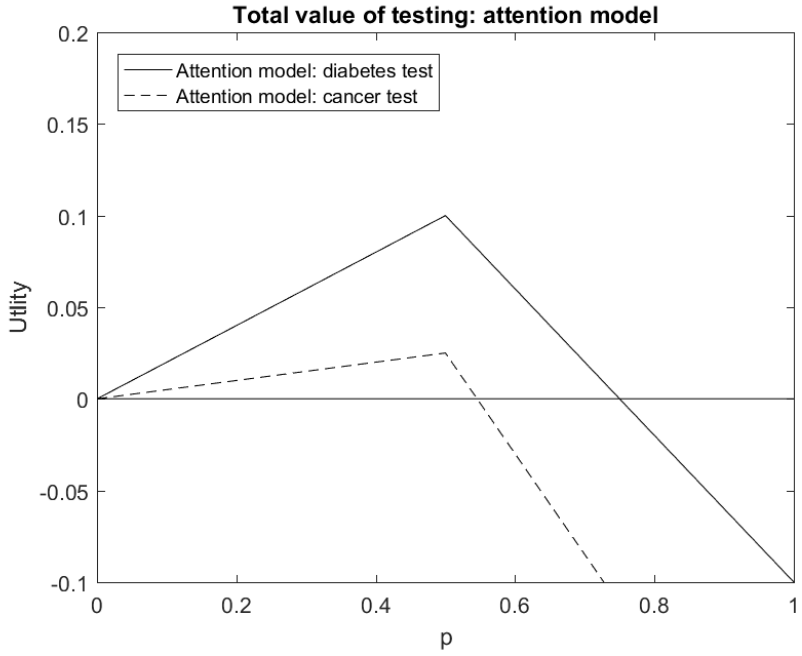
Note: Panel A illustrates predictions from the price experiment when we vary the cost of testing. For diabetes, we choose $\delta = 0.2,\ \Omega = 0.3,\ \Phi = 0.3,$ and $v = 0.$ The horizontal axis is p. The vertical axis is the total value of testing. C is the cost of testing. Panel B illustrates predictions from the disease experiment when the test is free. For cancer, we choose $v = 0.6$ to represent that individuals with more serious disease have lower health states if not treated.

Figure A4. Predictions of the Attention Model for the Price and Disease experiment

Panel A. Price experiment



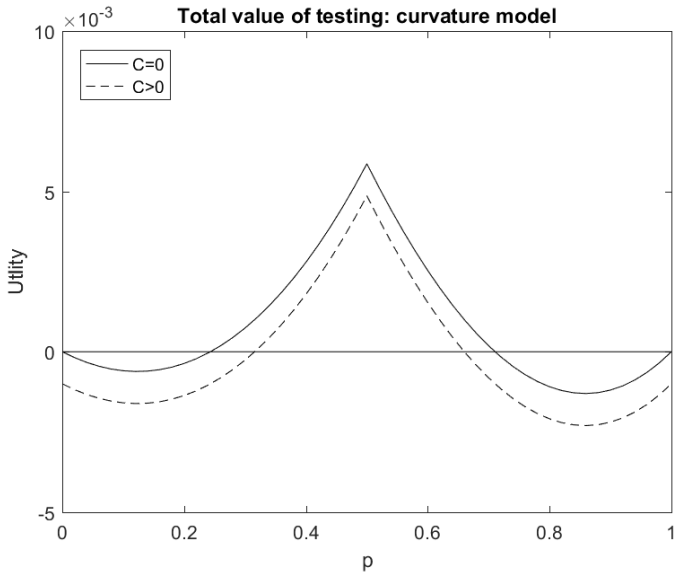
Panel B. Disease experiment



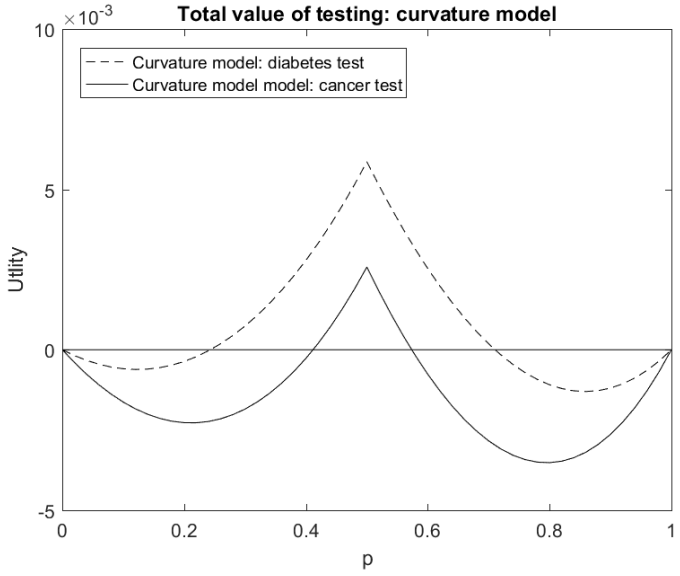
Note: Panel A illustrates predictions from the price experiment when we vary the cost of testing. For diabetes, we choose $a(\mu_0, \mu_1)$=0.1, $u_p = 1, \Omega = 0.3$, $\Phi = 0.3$. The horizontal axis is p. The vertical axis is the total value of testing. C is the cost of testing. Panel B illustrates predictions from the disease experiment when the test is free. We choose $a(\mu_0, \mu_1)$=0.25 for cancer to represent that individuals give more attention to more serious disease.

Figure A5. Predictions of the Curvature Model for the Price and Disease experiment

Panel A. Price experiment



Panel B. Disease experiment



Note: Panel A illustrates predictions from the price experiment when we vary the cost of testing. For diabetes, we choose $\alpha = 0.8,\ s_l = 1,\ s_h = 2$. When individuals do not take the test, $l(s,t^*) = \begin{cases} 0.05 & if\ \ s = s_l \\ 0 & if\ \ s = s_h \end{cases}$. The horizontal axis is p. The vertical axis is the total value of testing. C is the cost of testing. Panel B illustrates predictions from the disease experiment when the test is free. We choose $s_l = 0.9$ for cancer to represent that individuals with more serious disease have lower health states in low state.

Table A1. Constructing key variables in the summary statistics Table 1 from the survey questions

Panel A. Screening	
Take up rate of the test	=1 if take the test (from question 80)
Panel B. Demographics	
Gender (male)	=1 if male (from question 5)
Age	=2014-birthyear (from question 6)
Education years	(from question 8)
Marriage Status	=1 if married (from question 10)
Household Size	# of household members (from question 11)
Monthly income (⩾1000 RMB)	=1 if monthly income⩾1000 RMB (from question 18)
Panel C. Health Conditions and Behaviors	
Height (cm)	(from question 89)
Weight (kilogram)	(from question 90)
BMI ratio	=10000*weight/((height)^2)
Smoking (percentage)	=1 if smokes (from question 22)
Drinking (percentage)	=1 if drinks (from question 23)
Sleeping hours	(from question 24)
Exercise frequency	Larger if exercises less frequently (from question 44)
Subjective knowledge of diabetes	(from question 52), higher value means more knowledge
Objective knowledge of diabetes	=average of answers to question 52.1-question 57, higher value means more knowledge
Ability to follow treatment	=average of answers to question 58-question 60, higher value means more able to follow treatment requirements
Subjective assessment of disease risk	(from question 67), high value means more risk

Panel D. Preference Coefficients

Risk aversion	Report the switching point from (2) to (1). Larger switching point implies less risk aversion (from question 70)
Loss aversion	Report the switching point from (1) to (2). Larger switching point implies less loss aversion (from question 71)
Patience (includes present bias)	Report the switching point from (1) to (2). Larger switching point implies less patience (from question 72)
Patience (not includes present bias)	Report the switching point from (1) to (2). Larger switching point implies less patience (from question 73)
Monitoring	=average of answers to question 75.1-question 77.2, higher value means more tendency to pay attention to and avoid potential risk, equivalent to less information avoidance
Neuroticism	=average of answers to question 78.1-question 78.7 and question 78.9, higher value means more anxiety
Openness	(from question 78.8), higher value means more stable behavior and less risk taking

Table A2 . Robustness: the Non-monotonic Effect of Subjective Risk on Take-up Decisions

	(1)	(2)	(3)
VARIABLES	T0	T10	T30
2.new_percentile	-0.31**	-0.05	-0.05
	(0.16)	(0.13)	(0.10)
3.new_percentile	0.13	-0.02	0.15
	(0.11)	(0.12)	(0.11)
4.new_percentile	0.07	0.04	0.15
	(0.10)	(0.13)	(0.12)
5.new_percentile	0.23*	0.20	-0.08
	(0.12)	(0.19)	(0.10)
Constant	10.42**	2.98	-2.69
	(4.09)	(5.08)	(3.80)
Observations	204	197	209
R-squared	0.130	0.080	0.130
p-FTest2v3	0.015	0.882	0.179
p-FTest3v4	0.645	0.749	0.968
p-FTest4v5	0.276	0.460	0.108

Note: The regressions in the table show non-monotonic effect of subjective risk on take-up decisions in different treatment groups. We control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and ten variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. Robust standard errors in parentheses.

Table A3. Testing the Non-monotonic Effect of Subjective Risk on Take-up Decisions (Logit Model)

	(1)	(2)	(3)	(4)	(5)
	T0	T10	T30	Full Sample	Full Sample
Subjective risk	-0.47	-0.40	0.91**	0.37**	-0.48
	(0.66)	(0.49)	(0.45)	(0.16)	(0.61)
Subjective risk^2	1.66	0.95	-1.32*		1.70
	(1.29)	(0.69)	(0.71)		(1.25)
T10				-0.21***	-0.21***
				(0.04)	(0.05)
T30				-0.40***	-0.43***
				(0.04)	(0.05)
T10 × Subjective risk				-0.26	0.08
				(0.22)	(0.74)
T10 × Subjective risk^2					-0.86
					(1.40)
T30 × Subjective risk				-0.21	1.53*
				(0.23)	(0.81)
T30 × Subjective risk^2					-3.27**
					(1.52)
Demographics (6)	Yes	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes	Yes
Observations	204	197	196	610	610
R-squared	0.10	0.07	0.12	0.15	0.16

Note: The regressions test the effects of subjective risk on take-up decisions. Results are not affected by adding different categories of controls gradually. Marginal effects are reported but the significance level is indicated by the underlying estimates. Columns (1)-(3) estimate the effect separately in T0, T10 and T30, respectively. Columns (4) and (5) pool observations from T0, T10 and T30 together to test the cross-treatment difference. We control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and ten variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table A4. Testing the Effect of Subjective Risk on Take-up Decision: Disease experiment (Logit Model)

	(1) Diabetes treatment	(2) Cancer treatment	(3) Full sample	(4) Full sample
Subjective risk	0.81	0.70	0.60***	0.70
	(0.62)	(0.45)	(0.18)	(0.55)
Subjective risk square	-0.29	-1.21*		-0.25
	(1.22)	(0.70)		(1.06)
Cancer			-1.92	-1.51
			(3.70)	(3.57)
Cancer × Subjective risk			-0.53**	0.10
			(0.24)	(0.77)
Cancer × Subjective risk^2				-1.14
				(1.35)
Demographics (6)	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes
Observations	211	239	450	450
R-squared	0.14	0.14	0.14	0.14

Note: This table tests the effect of subjective risk on take-up decisions in disease experiment. Results are not affected by adding different categories of controls gradually. Marginal effects are reported but the significance level is indicated by the underlying estimates. Columns (1) and (2) estimate the effect separately in the diabetes and cancer treatments. Columns (3) and (4) pool observations from the two treatments together to test the cross-treatment difference. In the first two columns, we control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and ten variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. In the last two columns, we additionally control for the interaction terms of Cancer treatment with six demographic variables and ten variables on health conditions and behaviors, because the effects of these variables to the two diseases can be very different. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.

Table A5. Heterogeneous Effect in Disease experiment (Logit Model)

	(1) Diabetes treatment	(2) Cancer treatment	(3) Full sample	(4) Full sample
Subjective risk	0.87	12.07***	0.48**	0.73
	(0.73)	(3.20)	(0.19)	(0.63)
Subjective risk^2	-0.65	-20.43***		-0.55
	(1.32)	(5.35)		(1.12)
Cancer			-1.60	-0.75
			(4.04)	(3.85)
Cancer × Subjective risk			-0.78***	13.57***
			(0.28)	(2.52)
Cancer × Subjective risk^2				-23.68***
				(4.45)
Controllable	-0.14*	-0.07	-0.13**	-0.12*
	(0.07)	(0.06)	(0.06)	(0.06)
Controllable × Subjective risk	-0.04	-11.28***	0.34	-0.04
	(1.48)	(3.02)	(0.44)	(1.25)
Controllable × Subjective risk^2	1.08	19.19***		0.91
	(2.91)	(4.64)		(2.45)
Cancer × Controllable			0.03	0.03
			(0.10)	(0.10)
Cancer × Controllable × Subjective risk			0.16	-13.33***
			(0.52)	(2.75)
Cancer × Controllable × Subjective risk^2				21.85***
				(4.34)
Demographics (6)	Yes	Yes	Yes	Yes
Health Conditions and Behaviors (10)	Yes	Yes	Yes	Yes
Observations	176	207	383	383
R-squared	0.19	0.23	0.20	0.22

Note: This table tests the heterogenous effect of subjective risk on take-up decisions in disease experiment, depending on whether individuals believe the disease to be controllable or not. Results are not affected by adding different categories of controls gradually. Marginal effects are reported but the significance level is indicated by the underlying estimates. Columns (1) and (2) estimate the effect separately in the diabetes and cancer treatments. Columns (3) and (4) pool observations from the two treatments together to test the cross-treatment difference. In the first two columns, we control for six demographic variables—gender, age, education, marriage, household size, and monthly income—and ten variables on health conditions and behaviors: height, weight, BMI, sleeping hours, drinking behavior, smoking behavior, exercise frequency, subjective/objective knowledge of diabetes, and the ability to comply with diabetes treatment. In the last two columns, we additionally control for the interaction terms of Cancer treatment with six demographic variables and ten variables on health conditions and behaviors, because the effects of these variables to the two diseases can be very different. Robust standard errors in parentheses. ***, ** and * indicate significance at the 1%, 5%, 10% level, respectively.