Three Essays on Development Economics and Behavioral Economics

By

Changcheng Song

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Stefano DellaVigna, co-Chair
Professor Edward Miguel, co-Chair
Professor Matthew Rabin
Professor David Levine

Spring 2012

Abstract

Three Essays on Development Economics and Behavioral Economics

By

Changcheng Song

Doctor of Philosophy in Economics
University of California, Berkeley
Professor Stefano DellaVigna, co-Chair
Professor Edward Miguel, co-Chair

This dissertation studies retirement savings, weather insurance take-up and reference-dependent theory in the literature of development economics and behavioral economics. It consists of two field experiments and one laboratory experiment.

In Chapter one, I uses a field experiment to study the relationship between financial literacy and retirement savings in China. When the Chinese government launched a highly subsidized pension system in rural areas in 2009, 73% of households chose to save at a level that is lower than that implied by a benchmark life-cycle model. We test to what extent the low contribution level is due to a fundamental misunderstanding of the nature of compound interest. In a field experiment with more than 1000 Chinese households, we randomly assigned some households to a financial education treatment, emphasizing the concept of compound interest. This treatment increased the pension contribution by roughly 40%. The increase accounts for 51% of the gap between contribution levels in the Control group and those implied by the benchmark model. To pinpoint mechanisms, we elicited financial literacy after the intervention, and added a third group in which we explain the pension benefit in general. We find that the neglect of compound interest is correlated with low contributions to the pension plans in the control group, and that financial education about compound interest does help households partially correct their erroneous understanding of compound interest. Moreover, explaining compound interest increases their ability to translate benefits into their own situation. Welfare analysis suggests that financial education increases total welfare, although the fact that the treatment effects are heterogeneous implies that some households end up saving more than the level implied by the benchmark model.

In Chapter two (coauthored with Jing Cai), we use a novel experimental design to test the role of experience and information in insurance take-up in rural China, where weather insurance is a new and highly subsidized product. We randomly selected a group of poor households to play insurance games and find that it increases the actual

insurance take-up by roughly 48%. To pinpoint mechanisms, we test whether the result is due to: (1) changes in risk attitudes, (2) changes in the perceived probability of future disasters, (3) learning the objective benefits of insurance, or (4) the experience of hypothetical disaster. We show that the overall effect is unlikely to be fully explained by mechanisms (1) to (3), and that the experience acquired in playing the insurance game matters. To explain these findings, we develop a descriptive model in which agents give less weight to disasters and benefits which they experienced infrequently. Our estimation also suggests that experience acquired in the recent insurance game has a stronger effect on the actual insurance take-up than that of real disasters in the previous year, implying that learning from experience displays a strong recency effect.

In Chapter three, I conducted a controlled lab experiment to test to what extent expectations and the status quo determine the reference point. In the experiment, I explicitly manipulated stochastic expectations and exogenously varied expectations in different groups. In addition, I exogenously varied the time of receiving new information and tested whether individuals adjust their reference points to new information, and the speed of the adjustment. With this design, I jointly estimated the reference points and the preferences based on the reference points. I find that both expectations and the status quo influence the reference point but that expectations play a more important role. Structural estimation suggests that the model of the stochastic reference point fits my data better than that with expected utility certainty equivalent as the reference point. The result also suggests that subjects adjust reference points quickly, which further confirms the role of expectation as reference point

To my loving and supportive wife

Xiangge Wang

Acknowledgements

I am extremely grateful to Stefano DellaVigna, Edward Miguel, Matthew Rabin for encouragement, guidance, and helpful suggestions. I thank Liang Bai, Nicholas Barberis, Michael Carter, Keith Ericson, Frederico Finan, Benjamin Handel, Teck Ho, Jonas Hjort, Dwight Jaffee, Botond Kőszegi, Shachar Kariv, Ronald Lee, David I. Levine, Zachary Liscow, Ulrike Malmendier, Juanjuan Meng, Atif Mian, Edward Miguel, John Morgan, Terrance Odean, Matthew Rabin, Gautam Rao, Emmanuel Saez, Mike Urbancic and Jane Zhang for helpful comments and suggestions. I also thank the enumerators in China for their efforts in collecting the survey data, especially Jianing Yang for his help on facilitating the field work. I thank the People's Insurance Company of China for their close collaboration at all stages of the project. This research was supported financially by the NBER Household Finance Working Group Grant, Center for Equitable Growth, Xlab at UC Berkeley, and the UC Berkeley Program in Psychology and Economics (PIPE). The research also uses data supplied by the China Health and Nutrition Survey, funded by NIH (R01-HD30880, DK056350, and R01-HD38700) and the Carolina Population Center and the Chinese CDC.

## Chapter One: Financial Illiteracy and Pension Contributions-A Field Experiment on Compound Interest in China

### 1. Introduction

The inadequacy of retirement savings in the U.S. is a common, if not uncontroversial, theme in the literature.[1] Two bodies of literature have developed to explain this phenomenon. One, focusing on the lack of information and financial sophistication, stresses the importance of financial literacy and financial education. The other literature attributes under-saving to self-control problems and procrastination.[2]

This paper follows the first literature and uses a field experiment to study the relationship between financial illiteracy and retirement savings in China. We focus on one specific aspect of financial illiteracy, namely, the neglect of compound interest, and study whether financial education can improve people's understanding and change their behavior.

In China, although the savings rate is relatively high (Chamon and Prasad 2010), survey evidence suggests that rural households save little for their retirement due to the traditional reliance on children.[3] Yet, a dramatic fertility decline during the past few decades and increased longevity together are causing the population to age rapidly. Aging increases the burden on grown children to support their parents and challenges the tradition of saving little for retirement and relying on the children (Wang and Xia 1994; Wang 2000; Song 2001).[4] Population aging and the lack of retirement savings together cause social problems in rural areas, such as increasing tensions between the old and the young, and even spur rising suicides among old farmers (Zhang and Tang 2008). Therefore, the standard of living of the rural elderly has become an important concern for both researchers and policy makers.

In 2009 the Chinese government introduced the New Rural Social Pension Insurance Program (NRSPIP), which is voluntary and highly subsidized. Rural households can choose from a menu of five annual contribution levels: 100 RMB, 200 RMB, 300 RMB, 400 RMB, or 500 RMB, ranging from 2% to 8% of annual per capita net income in 2010.[5] The matching contributions from the government are: 30

---

[1] Diamond and Hausman (1984), Venti and Wise (1996), and Lusardi (1999) find that many households arrive at retirement with very little wealth. There is also opposite evidence: Scholz et al. (2006) find that most households in the Health and Retirement Study have accumulated more wealth than their optimal targets.

[2] This literature on financial literacy includes, but is not limited to, Lusardi (1999), Lusardi and Mitchell (2007b). The literature on procrastination includes Laibson et al. (1998), O'Donoghue an d Rabin (1999), Diamond and Kőszegi (2003), Choi et al. (2001), Madrian and Shea (2001).

[3] In a national survey of elderly, 10% of the rural elderly reported that they saved for their retirement and only 2% thought they saved enough (Guo and Chen 2009). In the China Health and Retirement Longitudinal Study, 4% of rural elderly reported that they relied on personal savings for old-age support, and 86% relied on their children (Zhao *et al.* 2009).

[4] By 2010, six working persons were supporting one retired person in China, but by 2050 fewer than two will support each retired person. I define "working persons" as those aged 15 to 60 and "retired person" as those aged 60 or over.

[5] 1 USD ≈ 6.35 RMB or 3.95 RMB in PPP; the annual per capita net income is around 6,500 RMB in

RMB, 30RMB, 40 RMB, 45RMB, and 50 RMB, respectively. The individual pension accounts consist of the individual contributions, the matching contributions, and the earned interest. Pensioners start to receive their pension at age 60, and the annual payout includes a share from individual pension accounts plus a 960 RMB subsidy. Given the high subsidies, the pension seems likely to be an attractive product prima facie.

Indeed, 93% of rural households in the study areas participated in the pension plans, but 88.5% of households contribute at the lowest level, 100 RMB. This is consistent with the survey evidence that rural households save little for their retirement (Guo and Chen 2009; Zhao *et al.* 2009). We show that a benchmark life-cycle model based on Gourinchas and Parker (2002) implies that 73% of households should save more in the pension plan than what we observe in practice, and they should increase their annual contribution by 80% on average.[6]The question, then, is why rural households do not save more for their retirement.

There are several possible explanations for the low level of retirement savings. Rural households might not trust that the government will deliver their pension in future. It is also possible that they save for retirement using other instruments, or plan to rely on their children. Although we cannot rule out these explanations, we will show some evidence in Section 8 that these are unlikely to be the main explanations for under-saving in our research setting.

In this paper, we explore another possible explanation: financial illiteracy. Research from the U.S. and other countries suggests that financial illiteracy is widespread and is correlated with poor decision making, even when the consequences are as significant as they are for retirement savings (Bernheim 1998; Lusardi and Mitchell 2007a, 2007b). The evidence on financial education is mixed and few can pinpoint the mechanism through which it works.[7] One possible mechanism is our focus here: there is evidence that individuals tend to linearize exponential functions when assessing them intuitively (Eisenstein and Hoch 2005; Stango and Zinman 2009; McKenzie and Liersch 2011).[8] For savings, such an error implies a systematic tendency to underestimate interest accrued in the future, in which case individuals will underestimate the value of saving. Stango and Zinman (2009) use the Survey of Consumer Finance to show that households with greater neglect of compound interest save less and borrow more. Yet there is little evidence on the causal effect of the neglect of compound interest on actual financial decisions.

We designed a field experiment to evaluate whether the neglect of compound interest is partially responsible for low level of contribution to pension plans in rural China. We randomly assigned more than 1000 Chinese households into three groups:

---

2010 in my study site (Municipal Bureau of Statistics 2011). This is 1,024 USD or 1,646 USD in PPP.

[6] In the benchmark model, we assume that they trust in the contract and there is no other channel to save for retirement except bank savings accounts. The details are discussed in Section 3.

[7] Some studies find small or no effects of financial education on individual decisions (Duflo and Saez 2003; Cole et al. 2011; Carter et al. 2008), while others find positive and significant effects (Bayer *et al.* 2008; Carlin and Robinson 2011; Gaurav et al. 2011; Cai and Song 2011).

[8] Stango and Zinman (2009) call the tendency to linearize exponential functions Exponential Growth Bias.

the *Control* group, the *Calculation* group, and the *Education* group. In the Control group, we visited households, explained the pension contract and did the survey. In the Calculation treatment, we calculated for the respondents the expected pension benefit levels after age 60 if they contributed at various levels with starting age 30. In the Education treatment, we asked them questions about compound interest, told them the correct answers, taught them the basic concept of compound interest, and did the calculation treatment. We then collected administrative data on their subsequent actual pension contributions.

To the best of our knowledge, this paper is the first to use a field experiment to identify the causal effect of the neglect of compound interest on real financial decisions.

We find that 56% of rural households in our sample were unable to provide a response to the simplest compound-interest question (after repeated prompting), and 73% of those who answered the question underestimated compound interest. Only 12% of rural households correctly estimated the compound interest or overestimated it. The result is similar to that in Lusardi and Mitchell (2007b), who find that only 18% of subjects in the Health and Retirement Study answered the compound interest question correctly.

Our experiment reveals that, although financial education had no effect on individual participation rates in the pension, it increased the annual contribution from 2 percentage points to 2.8 percentage points of annual per capita income, resulting in an increase of 40% relative to the average contribution of 133 RMB in the Control group. The increase accounts for 51% of the gap between the Control group's contribution and the level implied by the benchmark model, with a 95% bootstrapped confidence interval of 27% to 69%.

We then investigate the underlying mechanisms. We consider two possible explanations: learning the benefits of pensions in general, or better understanding of compound interest.

To assess the role of learning the level of pension benefits, we randomly assigned some households to a group in which we calculated for the respondents the expected pension benefit levels after age 60 if they contributed at various levels with starting age 30; we did not teach them about compound interest. We find that just doing the calculations and explaining the benefits increased the contribution by 20 to 25 RMB. This effect is significantly smaller than the treatment effect of education about compound interest discussed above. There might be two explanations: explaining why the benefit is large might increase the credibility of the described benefits, or increase the ability of translating the described benefits of age 30 into their own situation. We find that the treatment effects of education and calculation are similar for those who are around age 30, but differ when age increases. The treatment effect of calculation is lower than that of education for those who are around age 40, 50 or 60. Therefore, the different treatment effects between the education treatment and the calculation treatment are likely to be due to the ability to translate the benefit into their own situation.

To test whether this effect derives from a better understanding of compound interest, we measured financial literacy in the follow-up survey, and analyzed the relationship between the education intervention and financial literacy. We find that the neglect of compound interest is correlated with low contributions to the pension plans, and financial education on compound interest can help people improve their understanding. We then test whether education increases the understanding of compound interest to an extent that could generate the observed 53 RMB increase in contributions. We find that 34% to 81% of the treatment effects can be explained by a better understanding of compound interest, depending on the specification. The result suggests that understanding compound interest is a leading factor of the treatment effects, given the potential measurement error.

Welfare analysis shows that financial education increases total consumer welfare by 30% compared to the Control group, which is equivalent to a 3% increase in consumption each year after age 60. The welfare changes are heterogeneous: those who should save more do save more while some households end up saving more than the level implied by the benchmark model.

This paper contributes to the literature in the following ways. First, it helps to identify the barriers to the diffusion of new financial products, and contributes to the literature on technology adoption in developing countries. Financial products, such as pensions, can potentially help rural households smooth consumption, increase investment in human capital, and reduce poverty and vulnerability amongst the elderly.[9] The existing literature suggests that the use of these products is not widespread and provides evidence for a number of explanations (Gine *et al.* 2008; Cole *et al.* 2011). Yet the neglect of compound interest remains less explored as a possible explanation for the low utilization of savings products. We provide evidence that rural households in China underestimate compound interest and contribute less to pension plans.

Second, this paper adds to the existing evidence on the effect of financial education and identifies the mechanism through which it works. Although there is correlational evidence suggesting that individuals with low levels of financial literacy are less likely to participate in financial markets, plan for retirement, or transact in low-cost manners (Lusardi and Mitchell 2007a, 2007b; Lusardi and Tufano 2008), the experimental evidence on financial education is mixed. We provide more evidence that financial education can be effective in a real world intervention. More important, we show that just explaining the benefits of pensions is less effective than explaining the benefits plus providing specific education on compound interest. This suggests that teaching the underlying concepts can be particularly effective in changing behavior, which might be due to increases in the credibility of the described benefits.

Third, our results also contribute to the literature on consumer bias and pension savings. Existing literature suggests that many people do not save enough voluntarily to maximize their lifetime utility (Barr and Diamond 2008). Low savings for retirement can be driven by consumer biases, such as procrastination (Choi *et al.* 2001;

[9] For example, in Brazil, rural households containing pension receivers are less likely to experience income poverty than those without pension receivers (Barrientos et. al., 2003). In South Africa, the Old Age Pension program increases children's school attendance (Edmonds 2006) and improve their health and nutrition (Duflo 2000) because the pension is shared with them.

Madrian and Shea 2001). Neglect of compound interest is another plausible explanation for low savings that has not drawn much attention in the literature. If individuals neglect compound interest, they might underestimate the value of pension plans and thus contribute less than they should. This could lead to large welfare losses for them when they are older and have insufficient income. We build on previous studies that analyze the relationship between neglect of compound interest and saving decisions with laboratory experiments (Eisenstein and Hoch 2005) or observational data (Stango and Zinman 2009). Our approach goes beyond those studies by using a field experiment to identify a causal relationship between neglect of compound interest and actual saving decisions.

Furthermore, we show that we can improve consumers' financial decisions by correcting their erroneous understanding of compound interest. The psychology and economics literature has documented many individual biases. But whether these biases can be weakened is less explored. We build on the study of Eisenstein and Hoch (2005) and provide the evidence that we can debias the individual bias of neglecting compound interest.

Fourth, our paper adds to the growing literature that uses field experiments to test theory. We lay out a simple model of neglecting compound interest and test the qualitative implications of the model.[10]

The paper proceeds as follows. In Section 2, we provide background information on the rural pensions in China. In Section 3, we simulate the optimal level of pension savings. In Section 4, we describe the experimental design and survey data. The main empirical results are discussed in Section 5. In Section 6, we develop a simple model to explain the results. Welfare analysis is discussed in Section 7. Finally, we discuss alternative explanations in Section 8 and conclude in Section 9.

## 2. The New Rural Social Pension Insurance Program in China

China's population has been aging rapidly during the past few decades due to a fall in the population growth rate and an increase in life expectancy (see Appendix Table A1). By the year 2010, 12% of China's population was aged 60 years or over, and it is predicted that the number will increase to 34% by 2050 (United Nations 2011). Aging magnifies the burden on children to support their parents. Moreover, about 60% of the elderly people in China live in rural areas (State Council of the People's Republic of China 2006); they have accumulated relatively low incomes and savings during their working years. These facts cause many social problems in rural areas such as increasing tensions between the old and the young, and even suicides of old farmers (Zhang and Tang 2008; Sun Yefang Economic Science Fund Association 2010). Therefore, how to improve the standard of living of rural elderly has become a critical issue for the Chinese government, especially in recent years.

The New Rural Social Pension Insurance Program[11] was introduced in a few

---

[10] The literature on the role of theory in field experiments is reviewed in Card *et al.* (2011). Under their categorization, our experiment is a *Single Model* experiment.

[11] Before 2009, there were few alternative pension plans that were beneficial and affordable. There was the Rural Old-Age Pension Program, which was initiated in 1991 as an institutional framework for

pilot rural counties in 2009, and will expand to the whole country by the end of 2012. The new scheme is highly subsidized by the central and local governments. Farmers who are 16 years old or above, are not students, and are not enrolled in urban pension plans are eligible for the pension. The details of the plan are as follows. An individual lifetime bank account is established for the pension recipients. Each individual account of the pension fund is composed of individual contributions and government subsidies. Individuals can choose one of five annual contribution levels: 100 RMB, 200 RMB, 300 RMB, 400 RMB, or 500 RMB, which range from 2% to 8% of annual per capita net income in 2010. The Chinese government will provide subsidies to match the contribution according to Table 1, Panel A:

**Table 1. Pension Contract**

**Panel A: Pension subsidy**

| Options | Contribution level(RMB/year) | Government Subsidy(RMB/year) |
|---|---|---|
| 1 | 100 | 30 |
| 2 | 200 | 30 |
| 3 | 300 | 40 |
| 4 | 400 | 45 |
| 5 | 500 | 50 |

**Panel B: Example of Pension Benefit**

| | | | | | |
|---|---|---|---|---|---|
| Age when you start to contribute | 30 | | | | |
| Annual Contribution level | 100 | 200 | 300 | 400 | 500 |
| Annual Subsidy (RMB/year) | 30 | 30 | 40 | 45 | 50 |
| A: Basic pension after 60 years old | 960 | 960 | 960 | 960 | 960 |
| B: Amount from individual account balance (RMB/year) | 299 | 529 | 781 | 1023 | 1264 |
| C=A+B: Amount received annually after 60 years old (RMB/year) | 1259 | 1489 | 1741 | 1983 | 2224 |

Note: Panel A shows the corresponding government subsidy to each contribution level in the pension plans. Panel B provides an example to describe the explicit benefit of each contribution level for one who starts to contribute at age 30 and contribute for the next 15 years. The interest rate is assumed to be 2.5%, which is the one year interest rate in China at the time of this study.

Note that the marginal rate of subsidy decreases if individuals contribute more. All individual contributions and government subsidies will be deposited in the individual

---

administering a pension program based on voluntary-contribution, defined-contribution, and fully funded individual accounts (Shi 2006). The proportion of rural farmers insured under the program peaked in 1997 at 15.4%, but it declined to around 11% in 2004. The decline in the development of the rural old age security system was not only the result of mismanagement and the low coverage rate of the rural old age insurance system, but also stemmed from the government's unwillingness to make a financial commitment to set up such a system (Wang 2006). There were also pension plans offered by insurance companies, but they were too expensive and thus take-up was low.

account. The interest rate is the one-year base rate according to the People's Bank of China, the central bank, which is 2.5% as of 2011. The interest is compounded yearly.

Pensioners will receive their pension monthly after reaching age 60. The amount received includes two parts: a basic pension from the government and a portion from the individual account balance. The current basic pension is 80 RMB per month, or 960 RMB per year, which was 15% of per capita net income in 2010. The basic pension will be adjusted according to the price level of a given year. The amount paid out per month from individual accounts equals the individual account balance divided by 139 months. Therefore, the total amount received is:

$$\text{amount received per month} = \text{basic pension} + \frac{\text{individual account balance}}{139}$$

The new pension plans are highly subsidized by the central and local governments.[12] To illustrate, consider a farmer who is 30 years old and contributes the minimum amount (100 RMB) each year for 15 years. Assuming the interest rate is 2.5%, after age 60 the farmer is supposed to receive 1,259 RMB per year, of which about 82% comes from the government subsidy and its interest. If the farmer contributes 500 RMB, then approximately 56% would come from the government subsidy.

There are several special features of this pension program. For those who are already 60 or older, as long as all their eligible children living in the same village participate in the program, the parents can receive the basic pension every month without making any contributions. People between 45 and 60 years old are expected to contribute each year until they reach 60. Those under 45 years old should contribute each year for 15 years or more. Pension contributors may stop contributing for a few years and make up the contribution later. They can also cancel the pension and withdraw their savings. There is no subsidy if pensioners make up the contribution or cancel the pension. If pensioners die, their heirs will receive a lump sum payment that equals the remaining balance in the individual account minus the government subsidies.

## 3. Theoretical Framework

### 3.1 The Household Problem

To explain the pattern of pension savings, we apply a basic discrete-time, life-cycle model, augmented to incorporate uncertain lifetimes and uncertain incomes. We assume a finite horizon model in which individuals live to a maximum age $N$. Between ages 0 and $S-1$, individuals are children and make no consumption decisions. Adults start working at age $S$. At every age $S \le t \le T$, adults receive a stochastic income and decide how much to consume and how much to save for the future. Individuals stop working exogenously at the end of age $T$ and thereafter have no income if they do not participate in the pension program. There is one asset in the economy, with a constant interest rate $R$. We impose liquidity constraint so that illiquid assets cannot be borrowed against and liquid wealth must be weakly positive.

[12] In 2010, the contribution from farmers only accounted for about 25% of the total fund in my study county. The central government provides about 50% and the local government provides the other 25%.

Individuals also face a probability of death in each year of life. Individuals maximize their expected lifetime utility

$$\mathrm{E}[\mathrm{u}(\mathrm{C_S}) + \sum_{t=S+1}^{N} \beta^{t-S} u(C_t) \prod_{j=S}^{t-1} \mathrm{p_j}] \qquad (1)$$

Subject to

$$X_{t+1} = R(X_t - C_t - Q_t) + Y_{t+1} + Z_{t+1} \quad \textit{and} \quad X_{t+1} \geq 0$$

where $C_t$ represents total consumption at age *t*, $p_t$ is the probability that the individual at age *t* survives age *t+1*, $\beta$ is the discount factor, $X_t$ is cash on hand (total liquid wealth), $Q_t$ is the contribution to the pension at age *t*, $Y_{t+1}$ is the income at age *t+1* and $Z_{t+1}$ is the amount received from the pension fund after retirement.

The utility function is assumed to exhibit Constant Relative Risk Aversion:

$$u(C) = \frac{C^{1-\rho}}{1-\rho} \qquad (2)$$

To model the income uncertainty, we adopt Gourinchas and Parker's (2002) formulation, and decompose the labor income into a permanent component, $P_t$, and a transitory component, $U_t$;

$$\begin{aligned} Y_t &= P_t U_t \\ P_t &= G_t P_{t-1} N_t \end{aligned} \qquad (3)$$

The transitory shocks, $U_t$, are independently and identically log-normal distributed, $\ln U_t \sim N(0, \sigma_u^2)$. The log of the permanent component of income, $\ln P_t$, evolves as a random walk with age specific expected income growth, $\ln G_t$. The shocks to the permanent component of income, $N_t$, are independently and identically log-normal distributed, $\ln N_t \sim N(0, \sigma_n^2)$.

Note that there are some limitations in the benchmark model: we assume that the individual trusts in the contract and there is no other channel to save for retirement except bank savings accounts. In section 8, we will present some suggestive evidence that is consistent with these assumptions for our study sample.

### 3.2 Model Solution

Following Gourinchas and Parker (2002), we write the optimal consumption rule as a function of age, *t*, and normalized cash on hand, $x_t \equiv X_t / P_t$. The budget constraint becomes

$$x_{t+1} = (x_t - c_t - q_t)\frac{R}{G_{t+1}N_{t+1}} + U_{t+1} + z_{t+1} \qquad (4)$$

where lowercase letters are normalized by the permanent component of income. The Euler equation is:

$$u'(c_t(x_t)) = \beta R p_t E[u'(c_{t+1}(x_{t+1})G_{t+1}N_{t+1})] \qquad (5)$$

where $c_t(x_t)$ represents the optimal consumption rule at age t (normalized).

We estimate the real interest rate from return on Treasury bond and CPI. From 1981 to 2010, the average real interest is 2.26 percent, so $R$ = 1.0226. The number of patient options taken in Table A2 can be transformed to a range of discount factor $\beta$ in Appendix Table A3. Under the CRRA utility function, the number of riskless options taken in Table A2 can be transformed to a range of risk aversion parameters $\rho$ in Table A3. Both $\beta$ and $\rho$ are assumed to be the median of each range.[13]

We first use the China Health and Nutrition Survey (CHNS), to estimate income uncertainty and age-specific expected income growth. The CHNS is a longitudinal survey that includes eight waves, in 1989, 1991, 1993, 1997, 2000, 2004, 2006, and 2009. The survey covers coastal, middle, northeastern, and western provinces in China; see Appendix A for details.

We then solve the dynamic programming problem by solving the Euler equation for each choice of contribution level. We solve optimal consumption rules for each household based on age, time preference, and risk attitude. Then we simulate optimal consumption (and therefore wealth) each period for each household.

Finally, given the optimal life-cycle consumption path for each choice of contribution level, we can calculate the lifetime utility for each choice of contribution level and thus find the optimal contribution level in the rural pension program.[14] A complete description of the solution method is provided in Appendix A.

If the individual starts contribution at age $s$, the consumption at age $t$ will be

$$C_{t,s} = \begin{cases} C_t(X_t - q), if \quad s \le t < s+15 \quad and \quad t < 60 \\ C_t(X_t), if \quad s+15 \le t < 60 \\ C_t(X_t + \frac{12}{139} B_s(q) + 960), if \quad t \ge 60 \end{cases} \qquad (6)$$

Individuals are assumed to contribute the same amount for no more than 15 years before age 60. $G$ is the ratio between consumption after 60 and before 60. $B_s(q)$ is the individual account balance at 60 if the individual starts to contribute at age s and contributes $q$ for 15 years. Since the individual account balance will be distributed over 139 months, the amount received per year is $\frac{12}{139} B_s(q)$. The basic pension per year is 960 RMB. The individual account balance is calculated according to the pension contract:

$$B_s(q) = \sum_{t=s}^{s+14} (q + \tau(q)) \cdot (1+r)^{(60-t)} \qquad (7)$$

$\tau(q)$ is the subsidy for the contribution level $q$. $r$ is the one-year base rate from the

[13] Although there is no evidence that we can use elicited time and risk preference to calibrate lifecycle model, existing literature shows that they are correlated with actual economic outcomes (Tanaka *et al.* 2010). We do not intend to take our elicitations as accurate measure, but mainly to capture household variations. Sensitivity analysis shows that the calibration results are similar if we use $\beta = 0.96$ and $\rho = 0.5$

[14] In the simulation, we assume that people cannot change their contribution levels over time.

People's Bank of China.

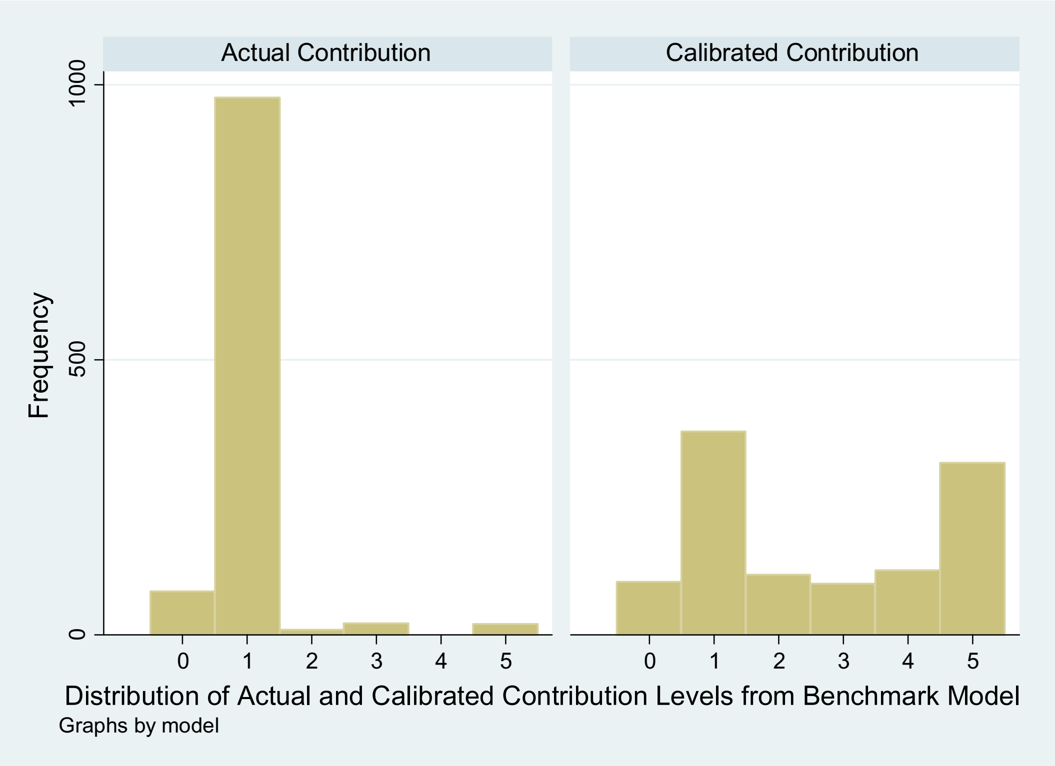



Figure 1 Distribution of Actual and Calibrated Contribution Levels from Benchmark Model

Note: The figures compare the distributions of the actual contribution and the calibrated contribution level from the model with correct perception of compound interest. The left figure shows the distribution of the actual contribution. The right figure shows the distribution of the calibrated contribution. The vertical axis is the density of the distribution. The horizontal axis is the contribution level from 1 to 5. The mean of actual contribution is 104 RMB and the mean of calibrated contribution is 234 RMB.

The above figure compares the distributions of the actual contribution and the calibrated contribution level. The left figure shows that around 90%of rural households chose the lowest contribution level. The right figure shows the prediction of the benchmark model. The benchmark model captures some aspects pretty well: most individuals participate in the pension. But the model captures other aspects poorly: individuals save more in the calibration than what we observe in practice.

We bootstrap the confidence interval of the calibrated contribution levels. To account for the correlation within each village, we use block bootstrap with each village as a block. The detailed procedure is discussed in Appendix A. We find the mean of the contribution level is 234 RMB, with a 95% confidence interval [213 RMB, 258 RMB]. The average actual contribution level is 104 RMB. Therefore, these calibration results suggest that rural households should save more in their pension plans.

If we try to use the benchmark model to explain the baseline contribution levels, one of the following three have to be true: (1) pensioners believe that the government or their grown children will give them 6000 RMB per year, which is roughly the annual per capita net income in 2010; (2) pensioners are extremely impatient with discount factors equal to 0.5; (3) pensioners believe that the government will deliver only 30% of their pension benefits. Therefore, the robustness checks suggest that the benchmark model is unlikely to explain the pattern of actual retirement savings. Rural households should save more in their pension plans.

### 4. Experimental Design and Survey Data

Our research site is in Shaanxi Province, whose economic development is around the mean of China, ranking 14th out of 34 provinces in 2009.[15] In 2011, 14 villages were randomly selected as experiment sites. The author, together with 14 hired enumerators who are college students, visited each village and conducted surveys of 1,153 households during the registration of new rural pension plans. Randomization of intervention was conducted at the household level. The timeline and intervention are presented in Figure 2 below.

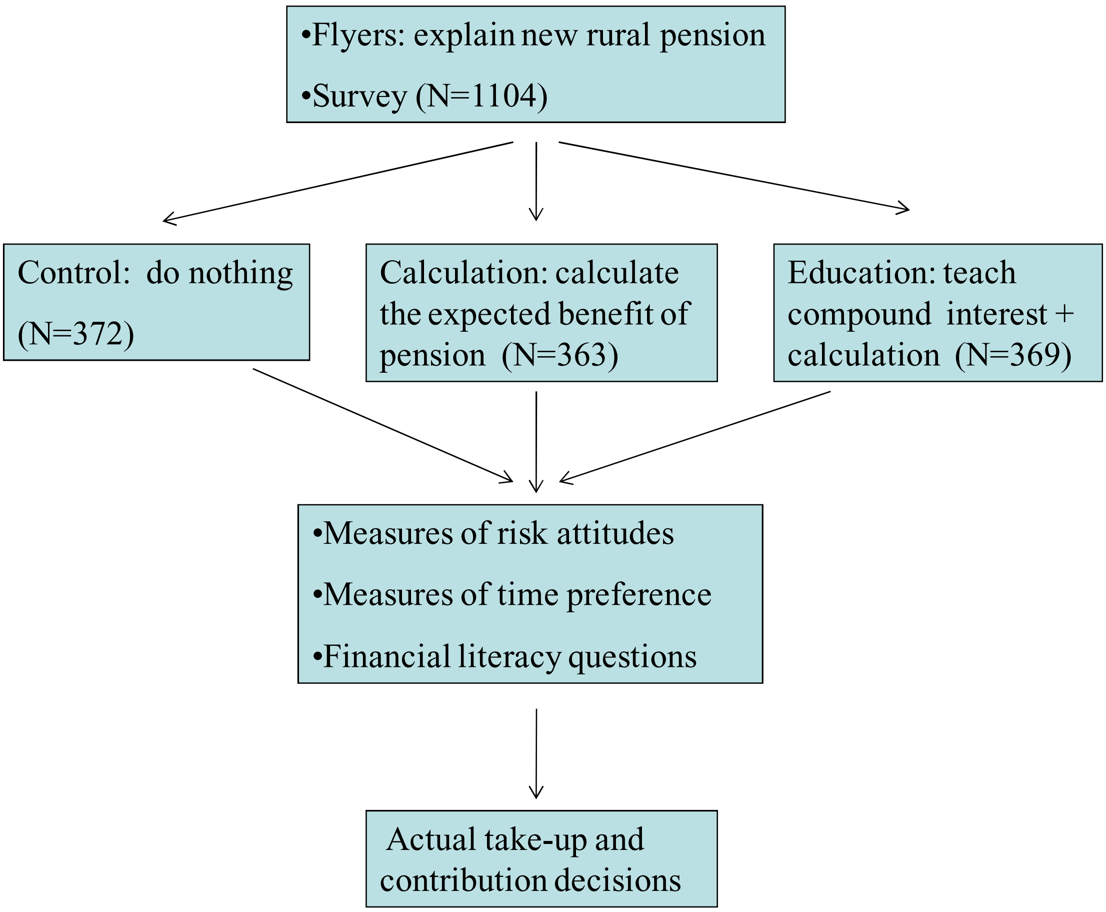



Figure 2 Timeline

During the household visits, the enumerators first gave households flyers with information about the new rural pension plans. We then asked households to fill out a

[15]Shaanxi Province is in the north-central part of China with two-thirds of its population from the rural area. By 2009, 12.8% of the rural population was aged 60 years or over (Municipal Bureau of Statistics 2011), which is slightly higher than the percentage for the whole nation. The income and consumption levels in this county are slightly higher than the national average of rural areas, ranging from 3% to 7% (Municipal Statistical Yearbook 2010; China Statistical Yearbook 2010).

survey about their socioeconomic background. Households were randomly assigned to three groups: the *Control* group, the *Calculation* group and the *Education* group (discussed below). For each group, we elicited risk attitudes, time preferences, and financial literacy (also discussed below). At the end of the visit, the enumerators asked sample households to indicate their contribution decisions. The decisions were passed to local village coordinators, who would collect the contributions later. We made clear that we were not employees of the government but independent researchers.

The details of the experiment are now discussed. In each village, households were randomly assigned to one of the three groups. In the *Control* group, the enumerators gave households the pension flyers and went over information about the contract. Then households were asked to fill out a short survey about their age, education, wealth, family members, risk attitudes, time preferences, and financial literacy.

In the *Calculation* group, the enumerators followed the same procedure but additionally calculated the expected pension benefits after age 60 if households were to contribute at various levels. The expected benefits are described in Table 1, Panel B. Enumerators went through the benefits of each contribution level with households and explained the range of differences. The purpose was to provide the explicit benefit amount of each contribution level *without* explaining the concept of compound interest. Comparing the Control group and the Calculation group will suggest whether explaining the benefits in general can increase the take-up and contribution level of pension plans.

In the *Education* group, the enumerators followed the same procedure as in the Control group and then asked questions about compound interest, taught the basic concept, and provided the calculated benefit for each contribution level. One key question about compound interest is adapted from Eisenstein and Hoch (2005):

*"You deposit 100 RMB as a Certificate of Deposit this year at a constant interest rate of 9% per year. Interest is compounded annually. How much money could you receive in 30 years?*

*1) Less than 300 2) 300-500 3) 500-1000 4) 1000-1500 5) More than 1500."*

No matter what participants' answers were, enumerators told them the right answer, 1,327 RMB, which is option 4. Then we briefly explained the basic concept of compound interest in a manner similar to Eisenstein and Hoch (2005): "Compound interest means that when interest is earned, it is left in the account. In future years, interest accumulates on the full amount that is in the account, so you earn interest on the interest as well as on the original principal amount." The other two questions are described in Table 3, Panel A. The purpose of this approach is first to document whether the farmers underestimate the value of savings from compound interest, and then to teach them about compound interest in order to debias them. Moreover, we also calculated expected benefits after age 60, as in the Calculation group.

To summarize, the *Calculation* treatment provides households with information about the expected benefits of each contribution level. The *Education* treatment makes households estimate interest, teaches the principle of compound interest, and provides households with information about the benefits.

Risk attitudes, time preferences, and financial literacy were elicited for all households. For those assigned to the Education group, the above three measures were elicited after education about compound interest. The comparison of these measures between the Education group and the other groups allows us to test whether education changes these parameters.[16] Risk attitudes were elicited by asking sample households to choose between increasing amounts of certain money (riskless option A) and risky gambles (risky option B) (see Appendix Table A2 Panel A).[17] We used the number of riskless options as a measurement of risk aversion.

Time preferences were elicited by asking sample households to choose between receiving some amount of money now (option A) and an increased amount of money one year later (option B) (see Appendix Table A2 Panel B). We used the number of patient options (option B) as a measurement of patience.

We also asked five questions to measure numeracy and financial literacy. These questions are described in Table 3 Panel B.[18] Note that Question 3 is similar to the compound-interest question in the education treatment.

---

[16] We did not ask the households the same questions before education, because households might be consistent within themselves so that we cannot see the treatment effects of these measures.

[17] Both time preference and risk attitude are elicited without money incentive.

[18] These questions were adopted from Banks et al. (2010), Lusardi and Mitchell (2006), Eisenstein and Hoch (2005), and Cole et al. (2011).

**Table 2. Summary Statistics**

| | Total | Control | Calculation Treatment | Education Treatment | p-value |
|---|---|---|---|---|---|
| **Panel A: Baseline** | | | | | |
| Male | 0.67 | 0.64 | 0.70 | 0.67 | 0.22 |
| | (0.47) | (0.48) | (0.46) | (0.47) | |
| Age | 44.90 | 44.87 | 44.40 | 45.42 | 0.30 |
| | (9.18) | (9.66) | (9.00) | (8.84) | |
| Years of schooling | 8.69 | 8.71 | 8.67 | 8.70 | 0.97 |
| | (2.50) | (2.56) | (2.56) | (2.40) | |
| Household size | 4.78 | 4.80 | 4.82 | 4.73 | 0.66 |
| | (1.34) | (1.37) | (1.38) | (1.29) | |
| Land for production | 3.75 | 3.75 | 3.76 | 3.73 | 0.98 |
| | (1.61) | (1.66) | (1.59) | (1.57) | |
| Share of agricultural income in total | 17.12 | 15.83 | 17.65 | 17.89 | 0.15 |
| | (16.64) | (14.17) | (17.30) | (18.18) | |
| Own business | 0.14 | 0.16 | 0.12 | 0.13 | 0.32 |
| | (0.34) | (0.36) | (0.32) | (0.34) | |
| Own a car | 0.10 | 0.13 | 0.11 | 0.06 | 0.0012 |
| | (0.30) | (0.34) | (0.31) | (0.24) | |
| Own a motorcycle | 0.44 | 0.48 | 0.44 | 0.40 | 0.09 |
| | (0.50) | (0.50) | (0.50) | (0.49) | |
| Saving for children | 0.81 | 0.79 | 0.83 | 0.80 | 0.50 |
| | (0.39) | (0.40) | (0.38) | (0.40) | |
| Saving for future when she/he is old | 0.25 | 0.26 | 0.26 | 0.25 | 0.92 |
| | (0.44) | (0.44) | (0.44) | (0.43) | |
| Number of children | 1.96 | 1.92 | 1.96 | 2.01 | 0.30 |
| | (0.84) | (0.90) | (0.80) | (0.80) | |
| Number of working children | 0.85 | 0.84 | 0.87 | 0.85 | 0.92 |
| | (1.03) | (1.01) | (1.05) | (1.03) | |
| Number of dependent old | 0.86 | 0.82 | 0.86 | 0.89 | 0.61 |
| | (0.89) | (0.87) | (0.91) | (0.89) | |
| Have a private pension plan | 0.13 | 0.14 | 0.13 | 0.13 | 0.83 |
| | (0.34) | (0.35) | (0.33) | (0.33) | |
| Take-up | 0.93 | 0.92 | 0.93 | 0.94 | 0.34 |
| | (0.26) | (0.28) | (0.26) | (0.23) | |
| Contribution level | 104.17 | 104.57 | 106.34 | 101.63 | 0.57 |
| | (65.28) | (71.23) | (70.03) | (53.14) | |
| **Panel B: Post-intervention** | | | | | |
| Risk aversion | 4.04 | 4.11 | 3.98 | 4.03 | 0.56 |
| | (1.68) | (1.65) | (1.71) | (1.69) | |
| Patience | 2.82 | 2.64 | 2.86 | 2.95 | 0.26 |
| | (2.61) | (2.64) | (2.61) | (2.59) | |
| Take-up | 0.98 | 0.98 | 0.98 | 0.99 | 0.83 |
| | (0.13) | (0.13) | (0.14) | (0.12) | |
| Contribution level | 157.16 | 133.06 | 156.19 | 182.38 | 0.00 |
| | (123.72) | (96.62) | (125.19) | (140.80) | |
| Observations | 1104 | 372 | 363 | 369 | |

Note: standard deviations are in the parentheses. P-values are for Wald test of equal means of three groups. *** significant at 1% level; ** significant at 5% level; * significant at 10% level

Table 2 presents summary statistics for the different groups. In total, we reached 1,330 households. A total of 177 households were not found, 32 households declined to participate in our study, and 17 households who were over 60 years old and cannot contribute to pension plans were mistakenly surveyed. Therefore, we have 1,104 surveys total. The overall attrition rate was 17.0%. The differences in attrition between groups are not statistically significant.

From Table 2, we see that the average education level of households is 8.69 years, which is close to graduation from secondary school. 13% of households have a private pension plan, which suggests that most households do not save for retirement in other pension plans. A total of 14% of households own a business, No household in my sample has any stock investment.[19] These results suggest that most households do not have other investments. Before our interventions, the take-up was 93% and the average contribution was 104 RMB (including those who did not participate), so most farmers participated in the pension plans but chose the lowest contribution level.

The last column shows the p-values for the Wald test of equal means of the three groups. Most control variables are balanced. The only exception is that the households in the Education group own fewer cars than those in the Control group and the Calculation group. However, the regressions in the next section show that the relationship between the contribution level and owning a car is in any case positive. Thus, this will not influence the validation of my randomization.

Table 3 presents the financial literacy of households. For different questions, the percentage of households that responded to the question and the percentage of households that answered it correctly vary. A total of 57.7% of households answered Question 4 correctly, which suggests that they have a basic understanding of inflation and purchasing power. A total of 13% of households answered Question 2 correctly and 5.6% of households answered Question 3 correctly, which suggests that most households have a poor understanding of compound interest.

[19] There is a concern that households do not like to report their investments and wealth. For the question of business ownership, most businesses are local and we actually visited their shops or factories to interview the owners. So it is unlikely that they lied to us. For questions about investment, given the financial knowledge rural people have, the misreporting is unlikely to be high.

**Table 3. Financial Literacy**

| | Question | Total | | Control | | Education Treatment | | Calculation Treatment | |
|---|---|---|---|---|---|---|---|---|---|
| | | %answer | %correct | %answer | %correct | %answer | %correct | %answer | %correct |
| | **Panel A: Questions used during the education treatment** | | | | | | | | |
| a | You deposit 100 RMB as a Certificate of Deposit this year at a constant interest rate of 9% per year. Interest is compounded annually. How much money could you receive in 30 years?<br>1) Less than 300 2) 300-500 3) 500-1000 4) 1000-1500 5) More than 1500 | | | | | 42.0 | 7.6 | | |
| b | Suppose you were 45 years old and you deposit 100 RMB every year for 15 years at a constant interest rate of 2.5% per year. Interest is compounded annually. How much could you withdraw when you are 60 years old?<br>1) Less than 1800 2) 1800-2000 3) 2000-2500 4)2500-3000 5) More than 3000 | | | | | 30.6 | 14.1 | | |
| c | Suppose you were 30 years old and you deposit 100 RMB every year for 15 years at a constant interest rate of 2.5% per year. Interest is compounded annually. How much could you withdraw when you are 60 years old?<br>1) Less than 1800 2) 1800-2000 3) 2000-2500 4)2500-3000 5) More than 3000 | | | | | 29.3 | 7.3 | | |
| | **Panel B: Post-intervention questions** | | | | | | | | |
| 1 | A second hand car is selling at 60000 RMB, which is 2/3 of the new one. What is the price of a new car?<br>1)90000 2) 40000 3)80000 4)120000 5)180000 6) other | 58.4 | 34.7 | 56.5 | 33.9 | 58.8 | 35.8 | 60.1 | 34.4 |
| 2 | If you borrowed 100000 RMB from the bank, the interest rate is 2% per month and compounded monthly. How much do you owe the bank in three months?<br>1) Less than 102000 2) 102000 3) 102000-106000 4) 106000 5) More than 106000 | 37.9 | 13.0 | 36.0 | 12.6 | 38.8 | 14.4 | 38.8 | 12.1 |
| 3 | You deposit 100 RMB as a Certificate of Deposit this year at a constant interest rate of 6% per year. Interest is compounded annually. How much money could you receive in 30 years?<br>1) Less than 300 2) 300-400 3) 400-500 4) 500-600 5) More than 600 | 33.4 | 5.6 | 29.8 | 3.5 | 35.0 | 7.0 | 35.5 | 6.3 |
| 4 | You deposited 10000 RMB in the bank and the interest rate is 2% per year. If the price level increases 3% per year, can you buy more than, less than, or the same amount of goods in 1 year as you could today? | 70.5 | 57.7 | 69.6 | 56.2 | 71.0 | 59.1 | 70.8 | 57.9 |
| 5 | You have two choices if you want to borrow 500000 RMB from the bank. Bank 1 requires you to pay back 600000 RMB in one month. Bank 2 requires you to pay back in one month 500000 RMB plus 15% interest. Which bank represents a better deal for you? | 52.5 | 22.8 | 49.7 | 22.0 | 53.1 | 24.1 | 54.8 | 22.3 |
| | Observations | 1104 | 1104 | 372 | 372 | 369 | 369 | 363 | 363 |

Note: The "%answer" equals the number of individuals who respond to the question divided by the number of observations in that column. The "%correct" equals the number of individuals who answer the question correctly divided by the number of observations in that column.

## 5. Empirical Results

In this section, we first document the fact that rural households underestimate the value of savings from compound interest. Then we show that financial education about compound interest can increase the households' contribution level. We also analyze the possible channels of the effects of financial education about compound interest.

### 5.1 Neglect of Compound Interest

We measure neglect of compound interest using the compound-interest question before intervention in the Education group. The response to the question is described in Figure 3A.

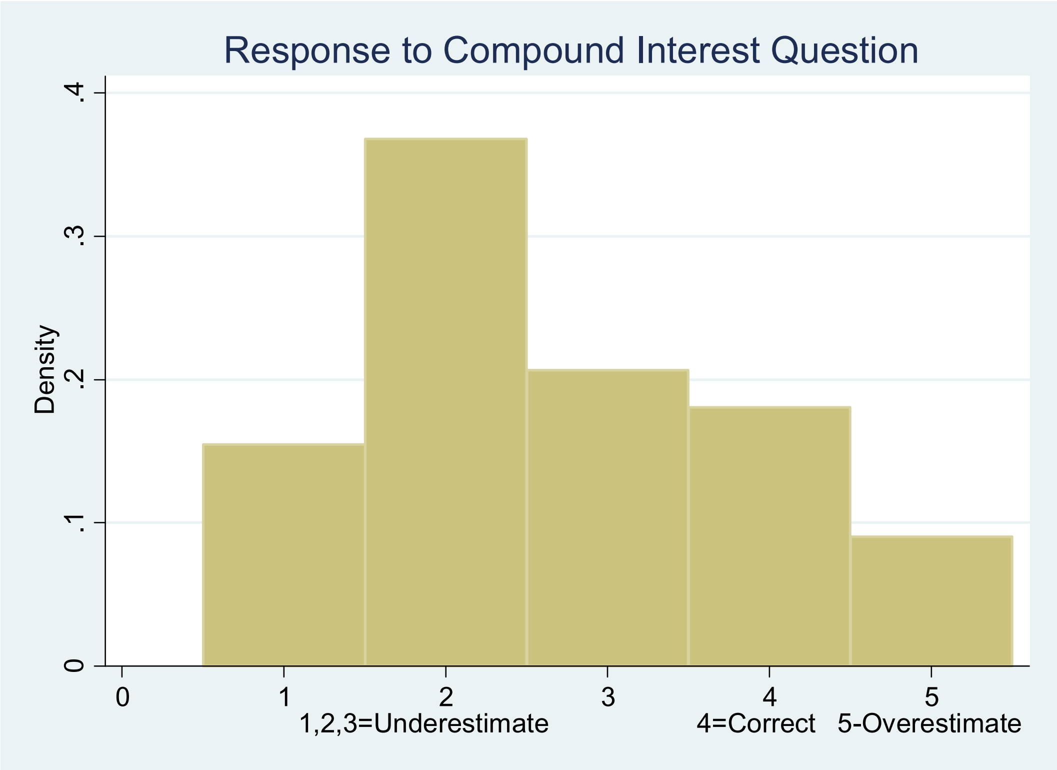



Figure 3A Response to Compound-Interest Question during the education treatment

Note: The figure shows the distribution of responses to compound interest rate question before intervention. The question is:"You deposit 100 RMB as a Certificate of Deposit this year at a constant interest rate of 9% per year. Interest is compounded annually. How much money could you receive in 30 years? 1) Less than 300 2) 300-500 3) 500-1000 4) 1000-1500 5) More than 1500"The figure only includes those who answered the question and excludes those who did not know.

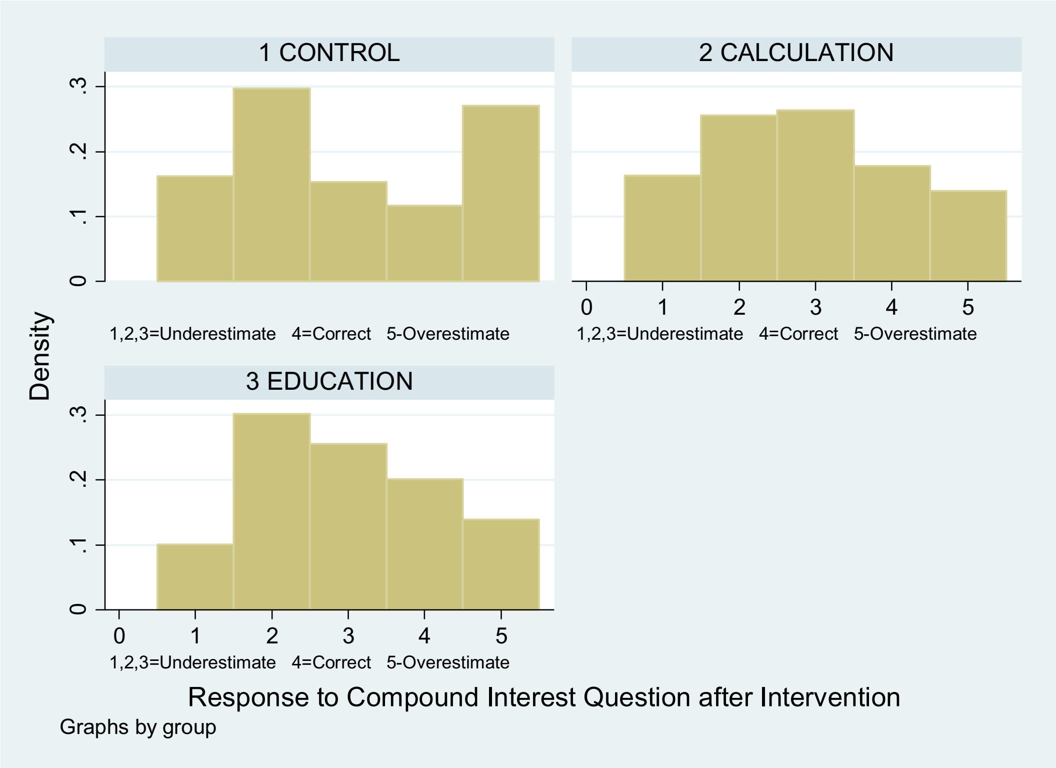



Figure 3B Response to compound-interest question after intervention

Note: The figure shows the distribution of responses to compound interest rate question after intervention. The question is:"You deposit 100 RMB as a Certificate of Deposit this year at a constant interest rate of 6% per year. Interest is compounded annually. How much money could you receive in 30 years? 1) Less than 300 2) 300-400 3) 400-500 4) 500-600 5) More than 600"The figure only includes those who answered the question and excludes those who did not know.

Out of 369 households in the Education group, 201 households were unable to provide the answer. Figure 3A only includes the 155 households that answered the question. The right answer is 1,327 RMB, which is option 4. A total of 18% of the 155 households chose the correct answer. 73% chose option 1 to 3, which can be characterized as underestimating the value from compound interest. And 9% chose option 5, which can be characterized as overestimating the value from compound interest. From Figure 3A, we can see clearly that rural households underestimate the value of savings from compound interest.

It is possible that households just randomly answered the compound-interest questions. In this case, the average should be 2.5 and the answer should distribute evenly across the five options. However, a t-test suggests that the average is different from 2.5 and it is significant at the 10% level. A chi-square goodness of fit test also rejects the hypothesis that the answers are uniformsly distributed at the 1% level. Therefore, it is unlikely that households just randomly answer the question; the evidence suggests that rural households underestimate the value of savings from compound interest.

### 5.2 The Impact of Education on the Take-up and the Contribution Level

Figure 4A shows that almost all the households in the three groups participate in the pension plans and there is no significant treatment effect. Figure 4B shows the treatment effect on the contributions. In the Control group, the average contribution is 133 RMB. In the

Calculation group, the average contribution increases to 156 RMB. In the Education group, the average contribution increases to 182 RMB. This suggests that both the education treatment and the calculation treatment increase the contribution level, and the education treatment is more effective.

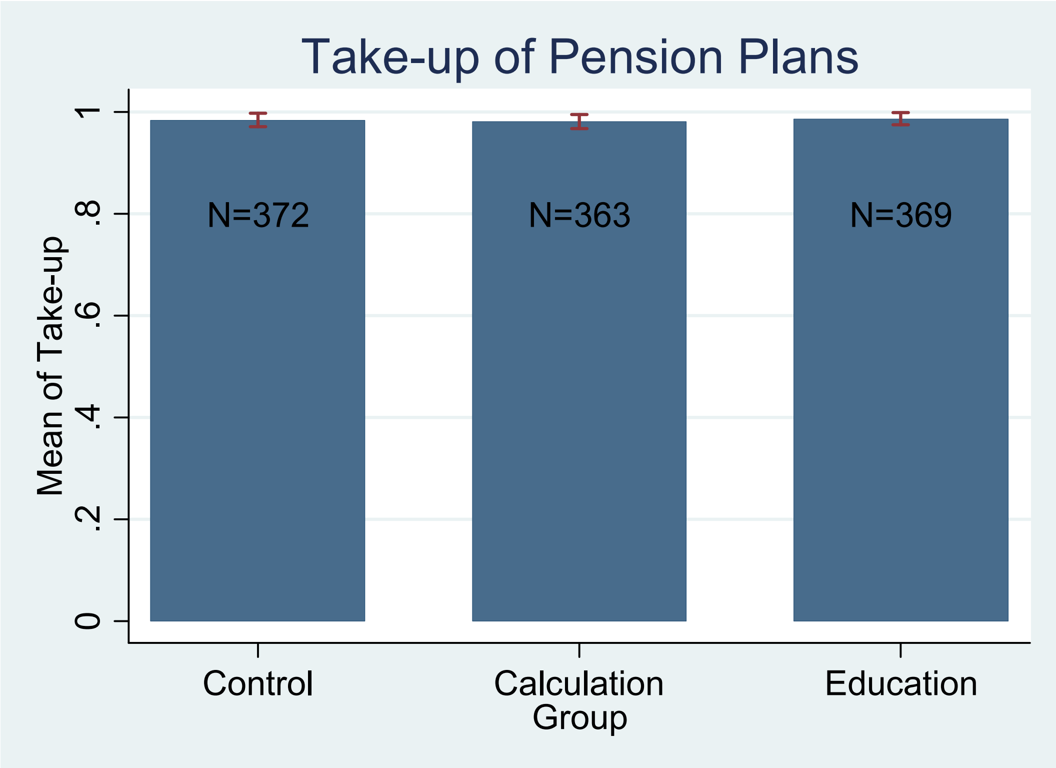



Figure 4A Treatment effect on Take-up

Note: This figure shows the treatment effect on the take-up of pension plans. In the Control group, the take-up is 98.4%. In the Calculation group, the take-up is 98.1%.In the Education group, the take-up is 98.6%. It suggests that almost all the households in the three groups participate in the pension plans.

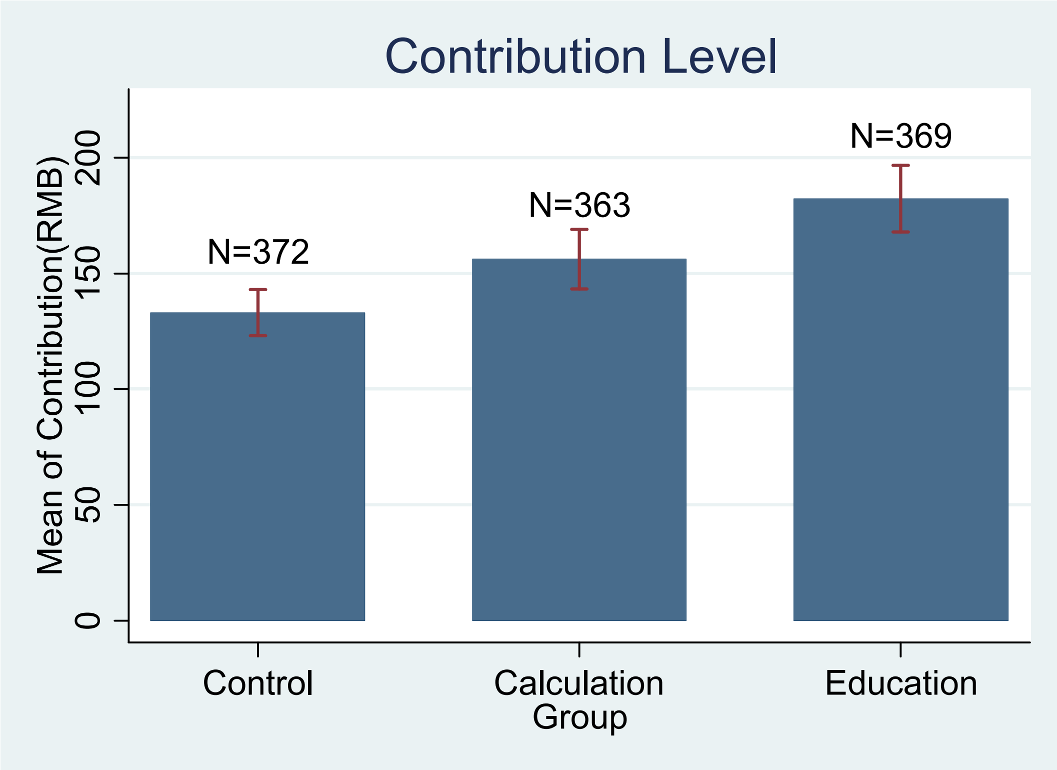



Figure 4B Treatment Effect on Contribution Levels

Note: This figure shows the treatment effect on the contributions of pension plans. In the Control group, the average contribution is 133 RMB. In the Calculation group, the average contribution increases to 156 RMB. In the Education group, the average contribution increases to 182 RMB. It suggests that both the education treatment and the calculation treatment increase the contribution level and the education treatment is more effective.

Figure 5 shows the distribution of contribution levels for different groups. Contribution level 1 is corresponding to 100 RMB contributions. Contributions level 2 to 5 are corresponding to 200 RMB to 500 RMB contributions, respectively. After the intervention, most individuals still contribute 100 RMB in the pension. In the Education group, there are more households contributing at 300 RMB and 500 RMB relative to the other two groups.

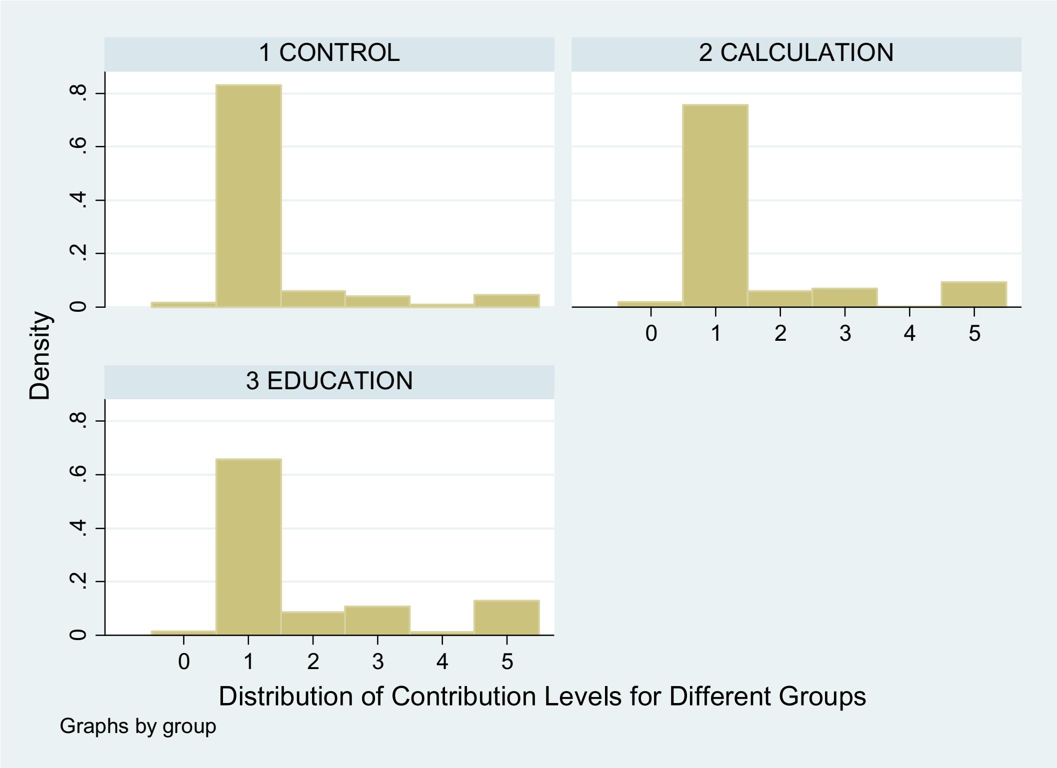



Figure 5 Distributions of Contribution Levels

Note: The figure shows the distribution of contribution levels for different groups. Contribution level 1 is corresponding to 100 RMB contributions. Contributions level 2 to 5 are corresponding to 200 RMB to 500 RMB contributions, respectively. It suggests that most individuals contribute 100 RMB and those in the Education group contribute more relative to the other two groups.

We estimate the treatment effect on the contributions through an OLS regression in (8):

$$q_{ij} = \alpha_j + \alpha_k + \beta_e \cdot Te_{ij} + \beta_c \cdot Tc_{ij} + \phi \cdot X_{ij} + \varepsilon_{ij} \qquad (8)$$

where $q_{ij}$ is the contribution levels or the changes of contribution levels for household i in natural village j. $Te_{ij}$ is an indicator for the education treatment and $Tc_{ij}$ is an indicator for the calculation treatment. Random assignment implies that $\beta_e$ is an unbiased estimate of the reduced-form intention-to-treat (ITT) education treatment effect and $\beta_c$ is an unbiased estimate of the ITT calculation treatment effect. $X_{ij}$ are household characteristics (e.g. gender, age, years of education, household size, land for production, car ownership, etc). $\alpha_j$ and $\alpha_k$ are village fixed effects and enumerator fixed effects, respectively. The covariates (X) and fixed effects are included to improve estimation precision and to account for chance differences between groups in the distribution of pre-random assignment (Kling, Liebman, and Katz 2007). The results are reported in Table 4.

**Table 4. The Effect of the Education and Calculation Interventions on Contribution Level**

| **Specification:** | **OLS regression** | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | **Individual Adoption of Pension** | | **Individual Contribution Level of Pension** | | **Change in Individual Contribution Level of Pension** | |
| **Sample:** | **All Sample** | | **All Sample** | | **All Sample** | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Education | 0.002 | 0.004 | 49.14 | 53.06 | 52.03 | 54.57 |
| | (0.009) | (0.009) | (9.39)*** | (9.28)*** | (9.27)*** | (8.94)*** |
| Calculation | -0.004 | -0.002 | 22.81 | 25.22 | 20.81 | 22.34 |
| | (0.009) | (0.009) | (9.16)** | (9.24)*** | (8.18)** | (8.13)*** |
| Male | | -0.011 | | 6.06 | | 10.93 |
| | | (0.006)* | | (10.18) | | (7.94) |
| Age (younger than 45) | | 0.0003 | | -1.83 | | -1.41 |
| | | (0.001) | | (0.79)** | | （0.078）* |
| Age (older than 45) | | 0.0008 | | 1.67 | | 0.69 |
| | | (0.001) | | (0.87)* | | （0.84） |
| Years of education | | 0.0003 | | 6.46 | | 5.82 |
| | | (0.002) | | (1.39)*** | | (1.42)*** |
| Household size | | -0.0006 | | -3.38 | | -5.14 |
| | | (0.004) | | (3.59) | | (2.91)* |
| Land for production | | 0.003 | | -1.68 | | -1.81 |
| | | (0.004) | | (3.90) | | (2.74) |
| Own a car | | 0.012 | | 26.39 | | 17.39 |
| | | (0.007) | | (16.00) | | (13.21) |
| Own a motorcycle | | 0.007 | | 15.30 | | 12.19 |
| | | (0.009) | | (8.38)* | | (8.11) |
| Wald test: $\beta_e=\beta_c$ | | | | | | |
| p-value | 0.4855 | 0.5009 | 0.0104** | 0.0064*** | 0.0007*** | 0.0004*** |
| Obs. | 1104 | 1104 | 1104 | 1104 | 1104 | 1104 |
| Omitted treatment | Control | | | | | |
| Mean of Dep. Var. for omitted treatment: | 0.0984 | | 133.06 | | 28.49 | |
| Fixed effects for village and enumerator | Y | Y | Y | Y | Y | Y |
| R-square | 0.0600 | 0.0648 | 0.0519 | 0.0895 | 0.0577 | 0.0963 |

Notes: Standard errors are clustered by 93 natural villages. Robust clustered standard errors are in the parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level. In columns 1 and 2, the dependent variable is individual contribution level. In columns 2 and 4, we add dummies for missing values of control variables in the regression. In columns 3 and 4, the dependent variable is changes in individual contribution level and we run the same regression as in column 1 and 2. $\beta_e$ is the coefficient of the education treatment and $\beta_c$ is the coefficient of the calculation treatment.

In columns 1 to 2, the dependent variable is individual take-up after our intervention. There is no evidence of treatment effect on take-up.

In columns 3 to 4, the dependent variable is the individual contribution level after our intervention. Column 3 presents results from the simplest possible specification, where the only right hand side variables are the indicators for the education treatment, the calculation treatment, and the fixed effects of natural villages and enumerators. The effect of the education treatment (49.14) is positive and significant at the 1% level. So the education treatment increases the contribution by 49 RMB; and it is around a 37% increase relative to the average contribution of 133 RMB in the Control group. The effect of the calculation treatment (22.81) is positive and it is statistically significant at the 5% level.

We calculate the degree to which these treatment effects can explain the gap between the Control group and the level implied by the benchmark model. We bootstrap the confidence interval of the percentage with a similar procedure in Section 3.2. We find that the treatment effect accounts for 51% of the gap between the Control group and the benchmark model prediction, with a 95% bootstrap confidence interval [27%, 69%].

In column 4, we add socioeconomic variables and dummies for missing values in the regression. The effects of the education treatment and the calculation treatment are similar to those in column 3. Years of education are positively correlated with the contribution level. Wealth, measured by owning a car or motorcycle, is also positively correlated with the contribution level.

In columns 5 to 6, the dependent variable is changes in individual contribution level, and we run the same regression as in column 3 to 4. Most coefficients have the similar magnitude and the same direction to those in the regression in which the dependent variables are individual contribution levels.

In sum, the education treatment increases the contribution by 49 to 53 RMB, resulting in an increase of around 37% to 40% relative to the average contribution of 133 RMB in the Control group. This suggests that our financial education has a positive and significant effect on retirement savings for rural households.

### 5.3 Possible Channels

In order for these findings to inform theory, more information is needed to analyze the mechanisms through which this effect could work. Possible explanations include: (1) learning the expected benefits of pensions in general, or (2) learning the expected benefits of pensions through better understanding of compound interest. The experiment is designed to be able to tell these mechanisms apart.

#### 5.3.1 Learning the Benefits of the Pension Program

It is possible that the education treatment provides direct information about the benefits of the pension. In Table 4, we find that the effect of the calculation treatment is positive and significant, which suggests that learning the benefits in general contributes to the overall effects. In order to test whether learning the benefits in general can fully explain the overall effect, we compare the treatment effect of the education treatment and the calculation treatment. The difference between those two interventions should indicate whether households acquire information about compound interest during the education. We report the p-value of the Wald test $\beta_e = \beta_c$ from Equation (8) in column 3 to column 6 in Table 4. The differences between $\beta_e$ and $\beta_c$ are between 26 RMB and 32 RMB. The impact of the education treatment is greater than the calculation treatment, and it is significant at the 1% level.

There might be two explanations for the difference between the education and the calculation treatment: explaining why the benefit is large might increase the credibility of the described benefits, or increase the ability of translating the described benefits of age 30 into their own situation.[20] Figure 6 shows the treatment effects of those two treatments for different ages.

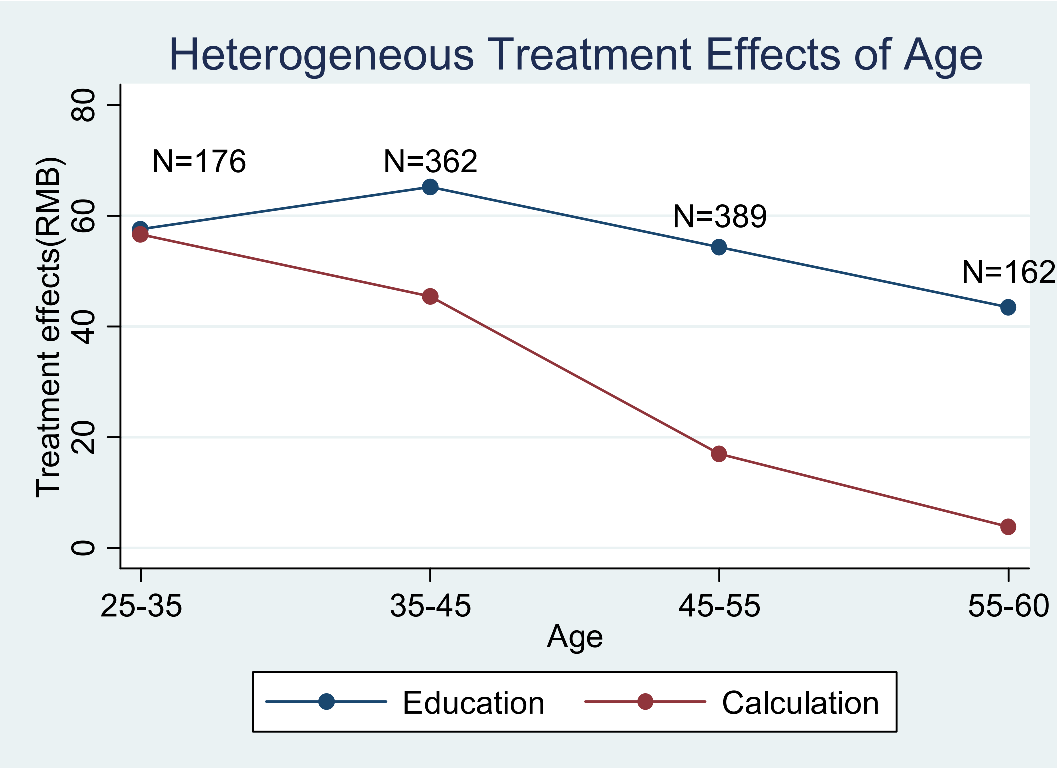



Figure 6 Heterogeneous Treatment Effects of Age

Note: The figure shows the heterogeneous treatment effects of age for the education treatment and the calculation treatment. The horizontal axis represents four age groups. The vertical axis is the treatment effects. The treatment effects of education and calculation are similar for those who are around age 30, but differ when age increases. The treatment effect of calculation is lower than that of education for those who are around age 40, 50 or 60. Therefore, the different treatment effects between the education treatment and the calculation treatment are likely to be due to the ability to translate the benefit into their own situation.

We find that the treatment effects of education and calculation are similar for those who are around age 30, but differ when age increases. The treatment effect of calculation is lower than that of education for those who are around age 40, 50 or 60. For those who are around age 30, the difference between the treatment effect of education and calculation is only 1 RMB. The differences are 19 RMB, 37 RMB and 39 RMB for those who are around age 40, 50 and 60. The difference is significant at the 5% level for those who are around age 50, but not significant for other age groups. Therefore, the different treatment effects between the education treatment and the calculation treatment are likely to be due to the ability to translate the benefit into their own situation.

### 5.3.2 Learning the Concept of Compound Interest

Another hypothesis is that individuals learn the concept of compound interest. Individuals may underestimate the value of savings from compound interest and thus contribute less to their pension plans. Financial education might increase household

[20] In the Calculation treatment, we calculated for the respondents the expected pension benefit levels after age 60 if they contributed at various levels with starting age 30. For those who are around 50, they need to infer their benefits by themselves.

contribution levels by helping households correct their erroneous understanding of compound interest. This hypothesis implies that the education treatment should correct households' erroneous answers to the compound-interest questions.

Figure 3B shows the response to the compound-interest question (Question 3) after intervention in different groups. Out of 1,104 households in the Education group, 725 households were unable to provide the answer. Figure 3B only includes the 369 households that answered the question, and excludes those that did not know. The right answer is 574 RMB, which is option 4. From Figure 3B, we can see clearly that rural households underestimate the value of savings from compound interest after intervention. A chi-square goodness of fit test also rejects the hypothesis that the answers are uniformly distributed at the 1% level. Therefore, it is unlikely that households just randomly answer the question, and the evidence suggests that rural households still underestimate the value of savings from compound interest.

Although neglect of compound interest still exists after intervention, there are fewer extremely wrong answers (option 1) and more correct answers (option 4) in the Education group than in the other groups. In order to take into account village fixed effects and other controls, we estimate the following equations:

$$q_{ij} = \alpha_{4j} + \alpha_{4k} + \beta_f \cdot F_{ij} + u_{ij} \qquad (11)$$

$$F_{ij} = \alpha_{5j} + \alpha_{5k} + \delta_e \cdot Te_{ij} + \delta_c \cdot Tc_{ij} + \phi_f \cdot X_{ij} + v_{ij} \qquad (12)$$

where $F_{ij}$ is the dependent variable measuring financial literacy. We use absolute distance to correct answer to measure financial literacy. Absolute distance measures how close the respondents' answers are to the correct ones. The absolute distance for each individual and each question is calculated in the following formula:

$$E(|x - x_c|) = \int_{x_l}^{x_u} |x - x_c| f(x) dx \qquad (13)$$

where $x$ is the chosen answer and $x_c$ is the correct answer. A complete description of the measurement is provided in Appendix B. Table 5 presents the estimation results in Equation (11) and (12).

**Table 5. The Effect of the Education and Calculation Interventions on Financial Literacy**

| **Specification:** | **OLS regression** | | **SUR regression** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dep. Var.:** | **Individual Contribution Level of Pension** | **Change in Individual Contribution Level of Pension** | **Absolute distance to the correct answer** | | | | | **Average standardized effect on absolute distance to the correct answer** | | |
| | | | **Question 1** | **Question 2** | **Question 3** | **Question 4** | **Question 5** | **Question 2 and 3** | **Question 1, 4 and 5** | **All questions** |
| **Sample:** | **Control** | | | | | | **All Sample** | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Education | | | 0.078 | 0.072 | 0.156 | 0.053 | 0.061 | 0.114 | 0.064 | 0.084 |
| | | | (0.067) | (0.068) | (0.071)** | (0.068) | (0.068) | (0.066)* | (0.049) | (0.050)* |
| Calculation | | | 0.072 | 0.024 | 0.121 | 0.029 | 0.043 | 0.072 | 0.048 | 0.058 |
| | | | (0.067) | (0.068) | (0.071)* | (0.068) | (0.068) | (0.058) | (0.046) | (0.047) |
| Absolute distance of Question 2 | -6.83 | 0.11 | | | | | | | | |
| | (9.65) | (5.37) | | | | | | | | |
| Absolute distance of Question 3 | 31.75 | 17.51 | | | | | | | | |
| | (8.71)*** | (5.93)*** | | | | | | | | |
| Obs. | 372 | 372 | 1104 | 1104 | 1104 | 1104 | 1104 | 1104 | 1104 | 1104 |
| Omitted treatment | | | | | | | Control | | | |
| Mean of Dep. Var. for omitted treatment: | | | -1.13 | -1.58 | -2.19 | -0.83 | -1.60 | | | |
| Social-economic variables | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Fixed effects for village and enumerator | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| R-square | 0.2013 | 0.2527 | 0.1752 | 0.1550 | 0.1632 | 0.1446 | 0.1545 | | | |

Notes: Standard errors are clustered by 93 natural villages. Robust clustered standard errors are in the parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level. In columns 1 and 2, we restrict the sample to the Control group and the Calculation group. In columns 3 to 7, the dependent variables are the absolute distance between the chosen answer and the correct answer for Question 1 to 5, normalized by standard deviation of Control group. In column 8, we report average standardized treatment effects on Question 2 and 3, of which both are compound-interest questions. The effect of financial education is positive and significant at the 10% level. In column 9, we report average standardized treatment effects on Question 1, 4 and 5, of which none of them are related to compound interest. The effect of financial education is positive but not significant. In column 10, we report average standardized treatment effects on all questions.

In columns 1 and 2, estimates from (11) are presented. Better understanding of compound interest is correlated with higher contributions. In columns 3 to 7, the dependent variables are the absolute distance between the chosen answer and the correct answer for Questions 1 to 5 (Questions 1 to 5 are described in Table 3) , normalized by standard deviation of the Control group. The effects of education on the financial literacy questions are all positive, but most are not significant. The only exception is Question 3, the compound-interest question. In column 5, the effect is positive and significant at the 5% level. So education reduces the distance from the correct answer by about one-sixth of a standard deviation. Therefore, financial education increases individuals' understanding about compound interest. Those in the Calculation group also have a better understanding of compound interest. It is likely that they infer large future benefits from the calculation treatment.

To illustrate the impact of the intervention on overall financial literacy, we follow Kling *et al.* (2004) and construct summary measures. Equation (14) defines average standardized treatment effects, $\tilde{\beta}$.

$$\tilde{\beta} = \frac{1}{\mathrm{K}} \sum_{k=1}^{K} \frac{\hat{\beta}_k}{\hat{\sigma}_k} \qquad (14)$$

where $\hat{\beta}_k$ is the point estimate for the treatment effect of outcome $k$ and $\hat{\sigma}_k$ is the Control group standard deviation of outcome $k$. To calculate the standard error for $\tilde{\beta}$, we need to account for the covariance of the estimates $\hat{\beta}_k$.We obtain this covariance matrix using the seemingly unrelated regression system shown in Equation (15).

$$Y = [I_K \otimes (T_e T_c X)]\theta + \nu \qquad (15)$$

where $I_K$ is a $K$ by $K$ identity matrix. The standard error and p-value for $\tilde{\beta}$ are based on the parameters, $\hat{\beta}_k$, jointly estimated as elements of $\theta$ in Equation (15).

In columns 8 to 10, we report average standardized treatment effects on three combinations of questions. In column 8, we report average standardized treatment effects on Questions 2 and 3, which are both compound-interest questions. The effect of financial education is positive and significant at the 10% level. In column 9, we report average standardized treatment effects on Questions 1, 4, and 5, none of which is related to compound interest. The effect of financial education is positive but not significant. In column 10, we report average standardized treatment effects on all questions, which is positive and significant at the 10% level. This suggests that financial education has a positive and significant effect on overall financial literacy, especially on the understanding of compound interest.[21]

To determine whether the education treatment increases understanding of compound interest and also increases the contribution level, we stack Equations (8), (11), and (12); generate indicators for each equation; and estimate the regression system following the same procedure in Section 5.3.1. We further replace Equation (11) with Equation (16), where we replace linear regression with quadratic functions because the relationship between the contribution level and understanding of

[21] Robustness checks suggest that other measures of financial literacy show similar results, such as squared distance, whether they answer the questions correctly and whether they answer the question.

compound interest is likely to be nonlinear.

$$q_{ij} = \alpha_{4j} + \alpha_{4k} + \beta_{f1} \cdot F_{ij} + \beta_{f2} \cdot F_{ij}^2 + u_{ij} \qquad (16)$$

We find that a better understanding of compound interest is unlikely to fully explain the main treatment effects. This might be due to a measurement error of financial literacy. A better understanding of compound interest can explain 7.4% of the treatment effects in the linear form and 33.8% of the treatment effects in the quadratic form. And they are both positive and significant at the 10% level.

We also run a 2SLS regression with Equation (12) as first stage and Equation (11) as second stage. We find that a better understanding of compound interest can explain 87% of the treatment effects in this specification.

To summarize, we find that although rural households underestimate the compound interest and contribute less to pension plans, education about compound interest can improve people's understanding of compound interest, and understanding compound interest is a leading factor of the treatment effects, given the potential measurement error.

**6. Models with Neglect of Compound Interest**

The evidence so far implies that education about compound interest can help to increase the contribution level by improving understanding of compound interest. In this section, we present a structural model to characterize neglect of compound interest, following Stango and Zinman (2009).

Consider an individual who saves an amount of money with present value (*PV*) at a periodic interest rate *i* over time horizon *t*, with periodic compounding. The future value (*FV*) is

$$FV = PV \cdot f(i,t) \qquad (17)$$

Following Stango and Zinman (2009), the term $f(i,t) = (1+r)^t$ is an exponential function, and an individual who neglects compound interest will underestimate $(1+r)^t$. Consider the individual who underestimates compound interest with the following form:

$$f(i,t,\theta) = (1+r)^{(1-\theta)t} \qquad (18)$$

$\theta$ measures the magnitude of neglect of compound interest: Unbiased consumers have $\theta = 0$ and correctly perceive compound interest, while those with $0 < \theta < 1$ neglect compound interest. Higher $\theta$ indicates greater neglect of compound interest.[22] Then perceived future values are calculated using

$$\mathrm{FV} = PV \cdot f(i,t,\theta) \qquad (19)$$

[22] This range of $\theta$ is relatively larger to that estimated by Stango and Zinman (2009), which is 0.2. This range of $\theta$ is relatively smaller to that estimated by Eisenstein and Hoch (2005) for savings, though they fit the slightly more flexible function $f(i,t,\alpha,\beta) = \alpha(1+r)^{\beta t}$ and estimate $\alpha = 0.35$ and $\beta = 0.36$.

If we incorporate neglect of compound interest into an intertemporal consumption model, the individual account balance is calculated in the following formula:

$$B_s(q) = \sum_{t=s}^{s+14} (q + \tau(q)) \cdot (1+r)^{(1-\theta)(60-t)} \qquad (20)$$

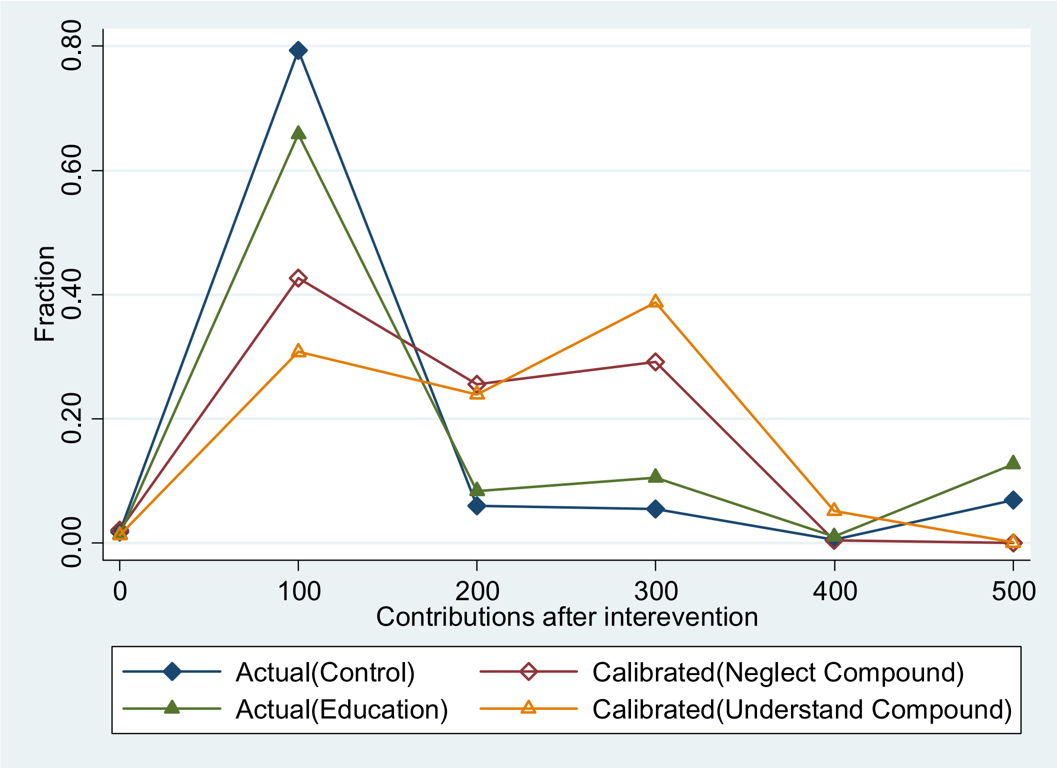



Figure 7 Distribution of Actual and Calibrated Contribution level

Note: The figures compare the distributions of the actual contribution in two groups, the calibrated contribution with neglect of compound interest and the calibrated contribution with correct perception of compound interest. The vertical axis is the fraction of each contribution. The horizontal axis is the contribution.

The above figure compares the distributions of the actual contribution in two groups, the calibrated contribution with neglect of compound interest, and the calibrated contribution with correct perception of compound interest. This shows that the calibrated contribution with neglect of compound interest can explain the change of the actual contribution in the Control and Calculation groups. This suggests that correction of erroneous understanding of compound interest can explain the effect of financial education about compound interest. Note the calibrated contribution with neglect of compound interest cannot fully explain the actual contribution in the Education group.

## 7. Welfare Analysis

### 7.1 Total Effects

In this section, we consider the welfare effect if households neglect compound interest based on the model in Section 6. We follow the framework of Liebman and Zeckhauser (2008). The basic idea is that if households correctly perceive compound interest, they should make the decision that maximizes their utility. However, if households neglect compound interest, they make their decisions to maximize their perceived utility ( $0 < \theta < 1$ ) but might make better decisions if they correctly perceive compound interest ( $\theta = 0$ ). The policy intervention of financial education should reduce their biases and thus help them make close to optimal decisions for their situation.

We use the benchmark model in Section 3 to calculate the welfare in each group. We find that the education treatment increased total consumer welfare by 30% compared to the Control group, which is equivalent to a 3% increase in consumption each year after age 60. This suggests that financial education increases total welfare.

### 7.2 The Distribution of the Effects: Targeting

A good policy intervention should increase total welfare of individuals. Ideally, policies should help people who behave suboptimally, but should have little negative impact on those who behave optimally (Camerer *et al.* 2003).

We check whether those households that should increase retirement savings really contribute more. We use our benchmark model in Section 3 to predict their contribution levels in the retirement plans. Then we divide the households into four groups: those who should not save more, those who should save 100 more, those who should save 200-300 more, and those who should save 400-500 more. We use Equation (1) to estimate the treatment effects $\beta_{nomore}$ , $\beta_{100}$ , $\beta_{200-300}$ and $\beta_{400-500}$ separately in these four groups, and compare the treatment effects. Figure 8 shows the heterogeneous treatment effects.

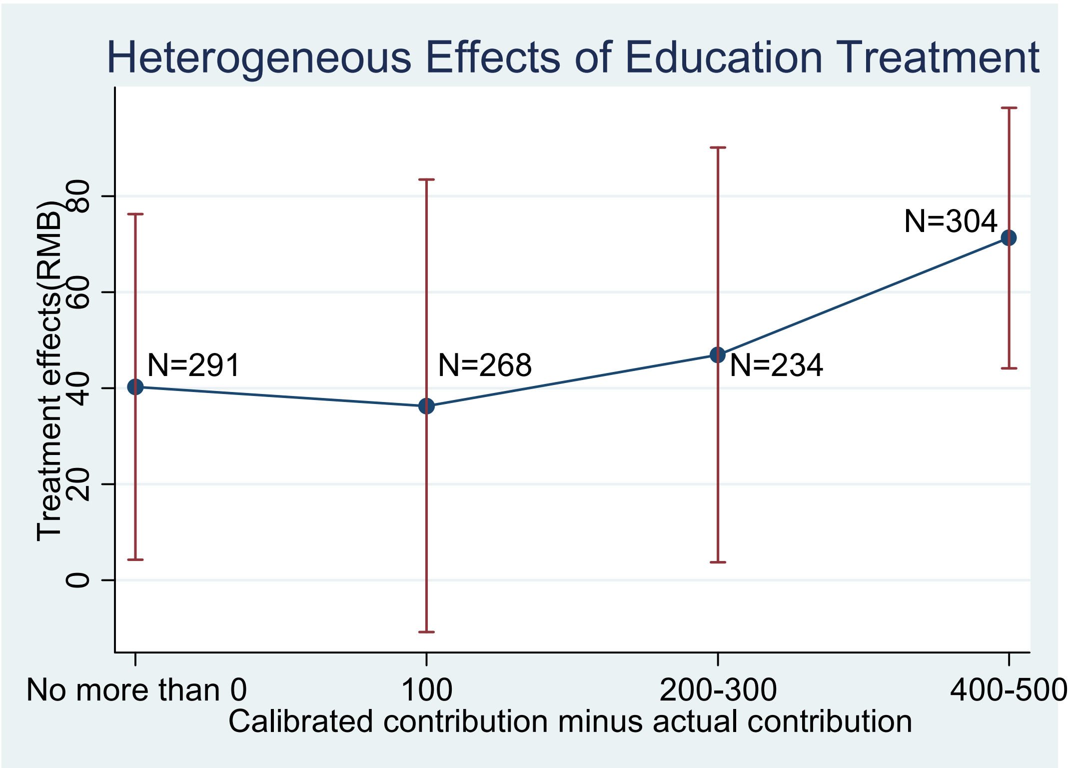



Figure 8 Heterogeneous Effects of the Education Treatment

Note: The figure shows the heterogeneous effects of the education treatment. The horizontal axis represents four groups based on the difference between the calibrated contribution and the actual contribution: those who should not save more, those who should save 100 more, those who should save 200-300 more, and those who should save 400-500 more. The vertical axis is the treatment effects.

The horizontal axis represents four groups based on the difference between the calibrated contribution and the actual contribution. The vertical axis is the treatment effects. We find that $\beta_{400-500} > \beta_{200-300} \approx \beta_{100} \approx \beta_{nomore} > 0$. Therefore, the welfare changes are heterogeneous: based on the benchmark model, those who should save more do save more while some households end up saving more than the level implied by the benchmark model.[23]

## 8. Alternative Explanations

There are some alternative explanations for why rural households save little in pensions. Although we cannot rule out these explanations, in this section we show evidence that they are unlikely to be the main explanations in my setting.

First, households might save for retirement in other ways. For example, they can save in private pension plans or invest in their own business. Although we cannot rule out this explanation, our survey suggests that this is not likely to be the case. Only 13% of households have a private pension plan, and only 14% of households own a business. Most households do not save for retirement via other sources.

Second, rural elderly might rely for old age on their children. China is a country which has a tradition of "rearing children for old age." In the China Health and Retirement Longitudinal Study, 86% of rural elderly reported that they relied on their children for old-age support (Zhao *et al.* 2009). However, population aging substantially increases the children's burden to support their parents. For example, by 2010, six working persons were supporting one old person in China, but fewer than two will support each old person by 2050.[24] Given China's rapid population aging, relying solely on children, without enough retirement savings might not suffice for living during old age.

Third, it is possible that rural households save little in pensions because they lack trust in the government. They might think that they will not receive the pension benefits when they are old. If so, financial education about compound interest should be less effective in the group with less trust in the government. In our survey, we asked about households' previous experience with the New Rural Co-operative Medical Care System, and use this to measure their trust.[25] For example, if people go to the hospital and do receive reimbursement from the government, they are likely to trust the government. If they go to the hospital but do not receive reimbursement, they are likely not to trust the government. We find that among people who have visited a hospital, only 8.6% did not receive reimbursement. Even in the group with lower

---

[23] There might be two reasons. First, there might be experimental demand effects so that all households save more in the pension. Second, there are some unrealistic assumptions in the benchmark model.

[24] I define "working persons" as those aged 15 to 60 and "old person" as those aged 60 or over.

[25] The New Rural Co-operative Medical Care System was a government program introduced in 2005 to overhaul the healthcare system in rural China. The annual cost of medical coverage is 50 RMB per person, of which 10 RMB is paid by the patient. The scheme will cover from 30% to 80% of their medical bill if patients go to a hospital.

measured trust, the treatment effect is positive and larger than the treatment effect in the whole sample but it is not significant due to small sample size. These findings suggest that trust in the government is unlikely to be the key reason for low pension savings.

Fourth, it is also possible that rural households save little in pensions because of liquidity constraints. If so, financial education about compound interest should be less effective in the group with less wealth. We use whether households own a business, and whether they own a car or motorcycle to measure their wealth. We find that even in the group with lowest measured wealth, the treatment effect is still positive, significant and close to the treatment effect in the whole sample. Moreover, the income per capita in my research site in 2010 was around 6,500 RMB (Municipal Bureau of Statistics 2011), of which the maximum contribution is less than 10%. And my benchmark model takes into account liquidity constraints. These findings suggest that liquidity constraints are unlikely to be the key reason for low pension savings.

Another alternative explanation is procrastination. Households might want to contribute more but procrastinate because of the immediate cost. However, there is no default, and everyone has to make a decision at a given time. Moreover, almost everyone participates in the pension plan but most only contribute 100 RMB. Therefore, procrastination is unlikely to be the key reason for low pension savings.

## 9. Conclusion

As rural households in developing countries tend to become old before they become rich, saving for retirement has become an increasingly important research and policy topic. Lack of pension savings can have significant consequences for the standard of living of the rural elderly. In this paper, we provide working age individuals with financial education about compound interest, and attempt to test for the role of neglect of compound interest in rural pension savings in China. We find that the education treatment increases contributions by 49 to 53 RMB, resulting in an increase of around 37% to 40% relative to the average contribution of 133 RMB in the Control group. We also investigate the possible mechanisms through which this effect might work, and find that learning the concept of compound interest is a primary factor.

Future research includes follow-up surveys of the pension and insurance programs to evaluate the long-term effects of financial education. Moreover, we will evaluate whether financial education influences households' behavior regrading other financial products. For example, theory predicts that better understanding of compound interest not only increases retirement savings but also other long-term savings.

The evidence on whether financial education can effectively change individual decisions is mixed, in the literature. This paper shows that learning the concept of compound interest can help to increase retirement savings in rural areas. Gaurav *et al.* (2011), and Cai and Song (2011) find that financial education with simulated experiences has a positive and significant effect on weather insurance adoption in developing countries. These findings suggest that we should first identify the barriers to individual participation and then apply specific financial education to remove the barriers. This seems to work better than general financial education.

From a policy perspective, this paper suggests that policy makers should take into account individuals' biases when designing policies, especially in rural areas where most people are poorly educated. In particular, policy makers can provide cheap financial education to overcome individual constraints, and thus improve individual welfare.

## Chapter Two: Insurance Take-up in Rural China-Learning from Hypothetical Experience

### 1. Introduction

Poor households in rural areas are vulnerable to losses from negative weather shocks (Banerjee 2003). To protect themselves from these shocks, they engage in costly ex ante risk-mitigation strategies, such as avoidance of high-risk and high-return agricultural activities, high levels of precautionary saving and insufficient investment in production (Rosenzweig *et al.* 1993) and human capital (Jesen 2000). The negative shock, the loss of profitable opportunities and the reduction of human capital accumulation can lead to persistent poverty.

A potential way to shield farmers from risks and to reduce poverty is to provide formal weather insurance products. In many cases, such insurance products are available but are not widely used.[26] In 2009, a rice insurance policy was first offered to rural households in Jiangxi Province of China. Under certain reasonable assumptions (discussed in Section 5), calibration suggests that more than 70 % of rural households should buy the weather insurance. However, the baseline take-up in our sample was only around 20%.These findings suggest a puzzle: why do so few households participate in weather insurance markets, given the potentially large benefit?

In this paper, we apply a novel method of financial education to test the role of experience and information in influencing weather insurance take-up, using a randomized experiment in rural China. Such insurance products are new to most farmers and large disasters are relatively uncommon.[27] Therefore, improving farmers' understanding of insurance benefits is important in this context.[28]

We offered financial education about weather insurance to a randomly selected group of households by playing insurance games with them. During the game, household heads were asked whether they would like to buy insurance for the hypothetical future year and then played a lottery to see whether there is disaster in that year. After the lottery results were revealed, the enumerator helped them to calculate the income from that year according to their insurance purchase decisions and the insurance contract. The game was played for 10 rounds. One or three days later, we visited sample households again to ask for their actual purchase decisions.

We find that playing insurance games increased the actual insurance take-up by 9.6 percentage points, a 48% increase relative to the baseline take-up of 20 percentage points. The effect is roughly equivalent to experiencing a 45 percentage point higher loss in yield in the previous year, or a 45 percentage point increase in the perceived

---

[26]For example, Gine, Townsend and Vickery (2008) find relatively low take-up (4.6%) of a standard rainfall insurance policy among farmers in rural India in 2004. Cole *et al.* (2008) also found relatively low take-up (5%-10%) of standard rainfall insurance in two regions of India in 2006. The take-up is higher (20%-30%) with door-to-door household visits.

[27] According to the private communication with local government officials, the actual probability of relatively large disaster in a year is around 10%.

[28] For example, in Gine et al. (2008), farmers who were asked why they did not buy weather insurance often responded that they "do not understand the product." This suggests that financial education might be important to help increase the use of insurance product.

probability of future disasters.

There are at least four possible mechanisms through which this effect could work: changes in risk attitude, changes in the perceived probability of future disasters, learning the benefits of insurance, and changes in experience of disasters and insurance benefits. We investigate each of them below.

After playing the insurance games, we elicited the subjects' risk attitudes and the perceived probability of future disasters. We then test whether playing insurance games increases either risk aversion or the perceived probability of future disasters by an amount that could generate the observed 9.6 percentage points increase in take-up. Our results show that it's not the case.

We also test whether this effect is due to learning the benefits of insurance by randomly assigning households to a group in which we explained the benefits of insurance. For these people, we calculated the payoff of the policy under different situations, but did not play insurance games. This treatment increases the actual take-up by only 2.7 percentage points, and the increase is not statistically significant. In fact, playing insurance games has a larger effect than just receiving the calculations, a difference which is significant at the 5% level. This suggests that learning the objective benefits of insurance is unlikely to fully explain the increased take-up.

To test whether this effect is driven by the experience of hypothetical disasters, we explore a second source of exogenous variation: the number of hypothetical disasters experienced during the game. We find that the total number of disaster increases take-up significantly and it is mainly driven by the number of disasters in last few rounds. Specifically, experiencing one more hypothetical disaster in the last five rounds increased the actual take-up by 6.7 percentage points. This suggests that the experience of recent disasters, even if hypothetical, might be the mechanism to influence the actual insurance decisions.

This paper contributes to the existing literature in the following ways. First, it sheds light on the puzzle of low weather insurance demand. Although existing research has tested a number of explanations (Gine *et al.* 2008; Cole *et al.* 2011), lack of experience remains less explored as a possible explanation. We provide evidence that the lack of experience of disasters and insurance contributes to the low take-up rate of weather insurance.

Second, this paper demonstrates a new method of financial education and shows that Although there is correlational evidence suggesting that individuals with low levels of financial literacy are less likely to participate in financial markets (Lusardi and Tufano 2008; Lusardi and Mitchell 2007; Stango and Zinman 2009), the experimental evidence of financial education is mixed.[29] We show that the novel method we used in this paper has a large and significant effect on improving insurance demand and it is more effective than the traditional method of financial education, which simply involves explaining the benefits.

Our results also contribute to the literature on the effect of direct experience.

[29] Some find small or no effects of financial education on individual decisions (Duflo and Saez 2003; Cole *et al.* 2011; Carter *et al.* 2008), while others find positive and significant effects (Cole *et al.* 2010; Gaurav *et al.* 2011; Cai, 2011).

Existing work has shown the effect of actual experience in areas including consumer behavior (Haselhuhn *et al.* 2009), financial markets (Choi *et al.* 2009; Agarwal *et al.* 2011; Malmendier and Nagel 2010) and charitable giving (Small *et al.* 2006). This paper analyzes the effect of hypothetical experience on poor households' insurance take-up and disentangles the effects of learning new information from the effects of personal experience. Results suggest that we can influence individual decisions by simulating experiences, as even hypothetical experience has an impact on household behaviors.

Fourth, this paper provides a new perspective on the role of laboratory experiments. Laboratory experiments provide controlled institutional contexts which are otherwise exceptionally difficult to obtain; they can generate deep insights about economic theories and policy applications (Holt 2005; Plott 2001). However, the behavior observed in the laboratory might not be a good indicator for behavior in the field under certain conditions (Levitt and List 2007). We demonstrate that laboratory experiments can serve as interventions in field experiments, by testing the causal effect of the laboratory experiment itself on actual behavior in the field. This differs from the more commonly used design of having all subjects participate in both a laboratory experiment and a field intervention, and correlating behaviors in the two (Ashraf *et al.* 2006; Gazzale *et al.* 2009; Fehr and Götte 2007). Unlike these studies, our random assignment procedure allows us to make a causal interpretation of the laboratory exposure. A difference from most laboratory experiments is that we paid all households a flat fee to eliminate confounding due to income effects.[30] It is interesting that, even when there is no incentive, we still observe a large treatment effect. Follow-up work will tell whether experiments with monetary incentives provide similar results.

The paper proceeds as follows. In section 2, we provide background information on rice insurance in China. In section 3, we describe the experimental design and survey data. The main empirical results are discussed in section 4. There, we present the main treatment effect of playing games on actual insurance take-up, analyze the possible channels of this effect and then show the dynamics of the take-up decision during the hypothetical games. Finally, in section 5, we develop a simple model to explain the results.

## 2. Rice Insurance in China

Nearly 50 percent of farmers in China produce rice, and rice is the staple crop for more than 60 percent of Chinese consumers. In 2009, The People's Insurance Company of China designed the first rice insurance program and offered it to rural households in 31 pilot counties. Our experimental sites are 16 natural villages within two rice production counties that were included in the first round pilots in Jiangxi

[30] The literature on financial incentives in experiments suggests that when there is no clear standard of performance in experiments, such as risky choices, incentives often cause subjects to move away from social desirable behavior toward more realistic choices (Camerer and Hogarth 1999). If social desirability depends on subject-experimenter interaction, households might buy more insurance during the games because of demand effects. In our data, the take-up during the games is around 75% and the actual take-up is around 27%.

province, which is one of China's major rice bowls.[31] All households in these villages were provided with the formal rice insurance product. Since the product was new at that time, no households had heard of or bought such insurance before.

The insurance contract is depicted in figure 1.

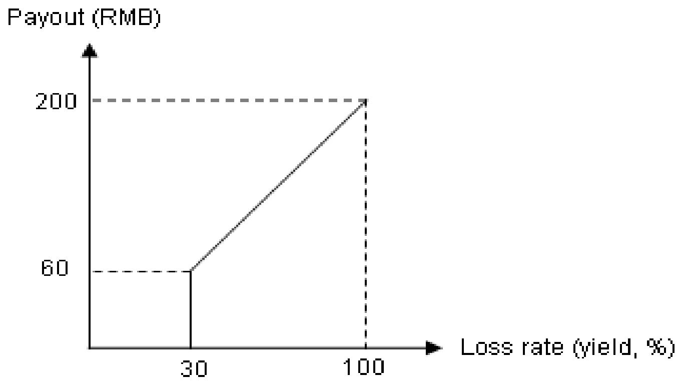



Figure 1 Insurance contract

Note: The original premium of insurance is RMB 12 per mu. The government will subsidize 70% of the premium so the households only pay the remaining 30%, i.e. RMB 3.6. The policyholder is eligible to receive a payment if there are disasters that cause 30% or more loss in yield by the following reasons: heavy rain, flood, windstorm, extremely high or low temperature and drought. The payout amount increases linearly with the size of the loss in yield, reaching a maximum payout at 200 RMB. The losses in yield will be determined by the investigation of a group of agricultural experts. They will come to the village to sample the rice in different plot and calculate the loss in yield.

The full insurance premium is 12 RMB per mu per season.[32]The government subsidizes 70 percent of the premium so that the households only pay 3.6 RMB. The policyholder is eligible to receive a payment if there are disasters that cause 30 percent or more loss in yield for one of the following reasons: heavy rain, floods, windstorms, extremely high or low temperatures, or drought. Losses in yield are determined by investigation by a group of insurance agents and agricultural experts. The payout amount increases linearly with the size of the loss in yield. For example, consider a farmer growing rice with an area of 2 mu. The normal yield per mu is 500kg but this year a wind disaster happens to reduce the yield to 300kg per mu. In that case, since the loss in yield is 40%, the farmer is supposed to get 200*40% = 80 RMB per mu from the insurance company. Note that the insurance is partial: payout is capped at 200 RMB, but the medium gross income in our sample is around 855 RMB per mu so the insurance covers at most 25 percent of income.

It's also important to note that the post-subsidy price is below the actuarially fair price according to our calculations. The profit of the insurance company is revenue minus payment to households and fixed cost.

$$\pi = N \cdot premium - N \cdot p \cdot indemnity - FC$$

where p is the probability of future disasters, N is the number of households who buy

[31] "Natural village" refers to the actual villages, "administrative village" refers to a bureaucratic entity that contains several natural villages.
[32] 1 USD≈6.35 RMB or 3.95 RMB in PPP; 1 mu≈666.7 $m^2$; 1 mu≈0.165 acre; Farmers produce two or three seasons of rice every year.

insurance and the indemnity is the payment to households when there is a disaster. According to private communications with local government officials, the actual probability of a disaster that leads to 30 percent or more loss is around 10 percent. Since $N \cdot 3.6 < N \cdot 10\% \cdot 60$ , the post subsidy price is below fair price. However, because the pre-subsidy price is higher than the fair price, the insurance company earns a profit if its fixed costs are not large.

## 3. Experimental Design and Survey Data

### 3.1 Experimental Design

In 2009 and 2010, we randomly selected 16 natural villages as our experiment sites. Nine hired enumerators consisting of government officials and primary school teachers, together with the two authors, visited each village and conducted surveys of 885 households before the beginning of the growing season. Randomization is conducted at the household level. There were two rounds of interviews for each household. The timeline is presented in the figure below.

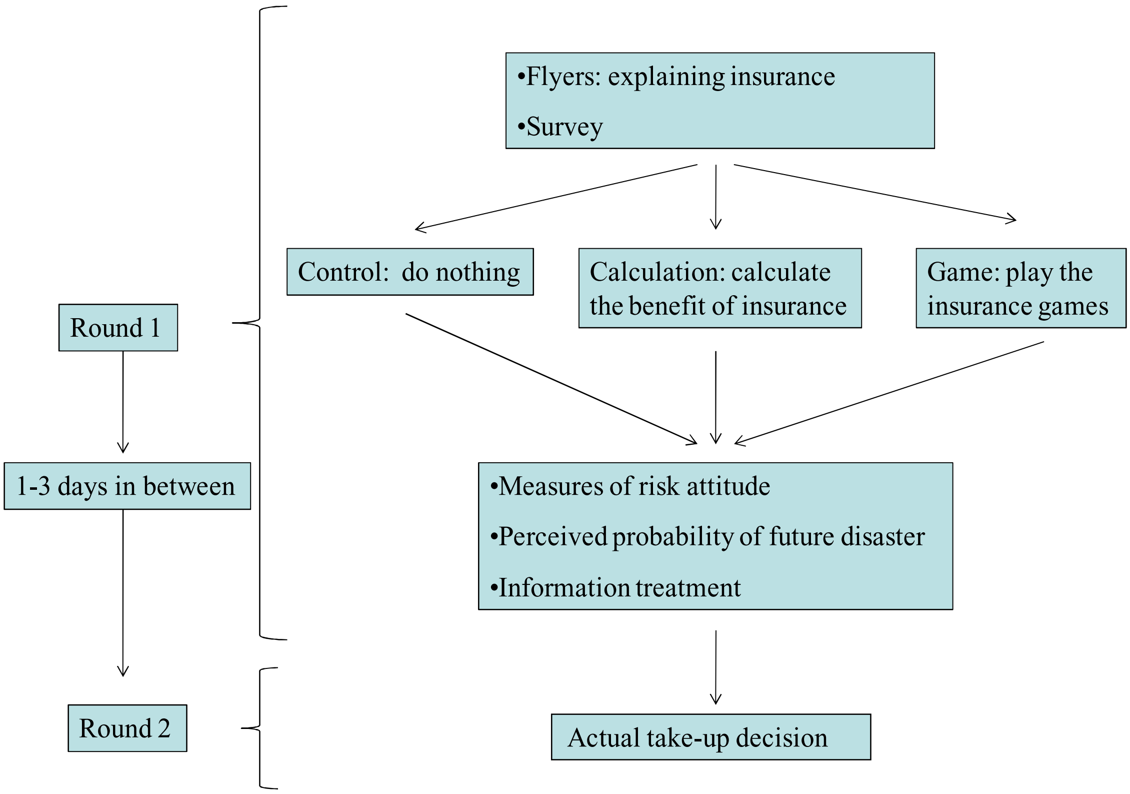



Figure 2 Timeline

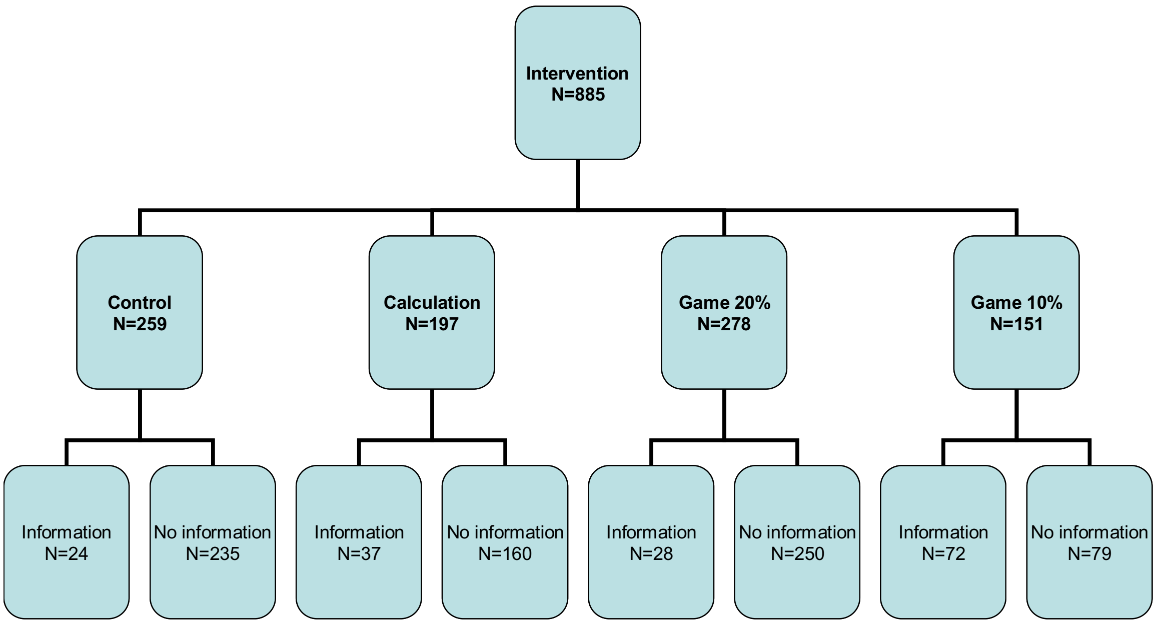
We implemented the baseline survey and intervention in round 1. The procedure is as follows: the enumerators first gave households flyers with information about the insurance contract, including liability, period and premium. Households were then asked questions about their socioeconomic background. If the households were assigned to the game treatment, the enumerators played the insurance games with them (discussed below). After the games, we elicited risk attitudes and the perceived probability of future disasters for all households (discussed below). If the households were assigned to the information treatment (discussed below), the enumerators informed them of the actual probability of a large disaster.[33] At the end of round 1, households were also told to think about whether they would like to buy the rice insurance and that enumerators would come back to ask them to make a decision in round 2.

Round 2 was conducted one to three days later. In round 2, the enumerators asked sample households to indicate their purchase decisions. The decisions would be passed to insurance company who would collect the premium later.[34]

Round 2 was conducted one to three days later. In round 2, the enumerators asked sample households to indicate their purchase decisions. The decisions would be passed on to the insurance company, which would collect the premium later.[35]

At the end of round 1, we paid each household 5 RMB to compensate for the participant's time. As discussed in the introduction, we did not incentivize decisions in order to eliminate confounding due to income effects.

We first approached the leaders of the villages and obtained a list that included the names of villagers and basic information about them[36].Then we stratified the households by their natural villages, ages of household heads, and area of rice production. In each stratum, households were randomly assigned to one of eight interventions. We randomized the treatments in two dimensions: how the contract was explained to the households (four groups) and whether the true disaster risk was revealed to the households (two groups). Figure 3 summarizes our design with eight groups in round 1.




[33] As estimated by government officials.

[34] Note that in round 1 the enumerators were randomly assigned to households while in round 2 one enumerator visited one or more villages. In our data, 22 percent of households (196 households) were visited twice by the same enumerator.

[35]Note that the enumerators were randomly assigned to households in round 1, while in round 2 one enumerator visited one or more villages. In our data, 22 percent of households (196 households) were visited twice by the same enumerator.

[36]We excluded households that did not grow rice. Those were households that were raising livestock or who had abandoned the land and were looking for jobs in urban areas.

The contract was explained in the following four ways. In the *Control* group, the enumerators gave households rice insurance flyers and went through the information about the contract. Then household heads were asked to fill out a short survey about their age, education, insurance experience, disasters experienced in recent years, production, social networks, risk attitudes and perception of the probability of future disasters.

In the *Calculation* group, the enumerators followed the same procedure as in the control group but additionally calculated the expected benefit of buying insurance if zero, one, two or three disasters were to happen in the following ten years. Enumerators went through the calculation with households and told them the summary: "According to our calculations, if there is no large disaster in next 10 years, it is better to not buy any insurance in the following 10 year. If there is at least one relatively large disaster, it is better to always buy insurance in the following 10 years."

In the *Game 20%* (and Game 10%) groups, the enumerators followed the same procedure as in the control group and then played the hypothetical insurance games with 20% (or 10%) probability of disaster for ten rounds. The game was played in the following way. Household heads were first asked whether they would hypothetically like to buy insurance in 2011 and then played a lottery with 20% (10%) probability of a disaster. We implemented the lottery by drawing randomly from a stack of cards; for example, in the *Game 20%*case, two out of ten cards signified disaster. After the lottery results were revealed, enumerators helped the household heads calculate the income from that year based on the expected income per acre and insurance payments. The game was then played for another nine rounds from hypothetical year 2012 to year 2020.[37]At the end of the game, we gave households the same information as in the *Calculation* group. Note that the game treatment provided not only financial education but also the second source of randomization: the number of the hypothetical disasters experienced during the games is randomized.

In a crossed randomization, we also randomized whether households were informed at the end of round 1 of the actual probability of disaster, which local government officials estimate at 10%. This randomization is interacted with how the contract is explained; thus, we have eight groups in total.

To summarize, the *Calculation* treatment provides households with information about the expected benefits of insurance. The *Game* treatment makes households acquire (hypothetical) disaster experience and provides households with information about the benefits of insurance. The (crossed) *Information* treatment provides households with information about the risk of disaster.

Risk attitudes and the perceived probability of future disasters were elicited for all households. For those who were assigned to play games, the above two measures were elicited after playing the insurance games. Comparing these measures between the game group and the other groups allows us to test whether playing games changes

---

[37]The setup implies that 89 percent of households in the *Game 20%* group and 65 percent of the households in the *Game 10%* group were expected to experience at least one disaster. In our data, 82 percent of households in the *Game 20%* group and 66 percent of households in the *Game 10%* group experienced at least one disaster.

these parameters and further changes the actual insurance take-up.[38]Risk attitudes were elicited by asking households to choose between increasing amounts of certain money (riskless option A) and risky gambles (risky option B) in Appendix TableA2 Panel A. We use the number of riskless options as a measurement of risk aversion.

The perceived probability of future disasters was elicited by asking households "what do you think is the probability of a disaster that leads to more than 30 percent yield loss next year?" We used a simple mechanism to illustrate probability, which might be a difficult concept for households with limited education.[39]

### 3.2 Survey Data

We implemented the survey in three waves. In the first wave (181 households, August 2009), we implemented only the control and *Game 20%*in the no information treatment. In the second wave (379 households, early March 2010), we implemented the control, the *Calculation*, and *Game 20%*in the no information treatment. In the third wave (325 households, late March 2010), we implemented all eight interventions. Because the *Game 10%* group and the information treatment were only conducted in the third wave, we oversample the *Game 10%* group; the total sample sizes of the *Game 10%* group and the information treatment are smaller than in the other groups.

**Table 1. Summary Statistics and Randomization Check**

| | Wave 1 | | | Wave 2 | | | | Wave 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Control | Game 20% | p-value* | Control | Calculation | Game 20% | p-value** | Control | Calculation | Game 20% | Game 10% | p-value** |
| **Panel A:before playing the game** | | | | | | | | | | | | |
| Age | 46.90 | 50.44 | 0.05 | 51.43 | 50.86 | 52.99 | 0.34 | 50.64 | 48.27 | 52.10 | 48.53 | 0.23 |
| | (11.33) | (12.37) | | (11.41) | (11.67) | (12.32) | | (12.28) | (11.47) | (12.24) | (12.17) | |
| Education*** | 1.38 | 1.32 | 0.57 | 1.30 | 1.30 | 1.35 | 0.84 | 1.45 | 1.37 | 1.41 | 1.44 | 0.94 |
| | (0.75) | (0.82) | | (0.78) | (0.71) | (0.82) | | (0.78) | (0.85) | 0.93) | (0.90) | |
| Household Size | 4.80 | 5.04 | 0.62 | 5.05 | 5.25 | 5.26 | 0.80 | 4.48 | 4.60 | 4.31 | 4.58 | 0.75 |
| | (1.79) | (2.30) | | (2.52) | (2.84) | (2.89) | | (1.29) | (1.39) | (1.69) | (1.51) | |
| Area of Rice Production (mu) | 12.14 | 12.08 | 0.97 | 8.90 | 9.20 | 8.90 | 0.94 | 10.28 | 11.91 | 10.46 | 11.25 | 0.69 |
| | (9.58) | (7.56) | | (7.51) | (7.90) | (7.79) | | (5.42) | (13.57) | (10.25) | (7.37) | |
| Share of Rice Income in Total Income (%) | 84.00 | 85.05 | 0.76 | 64.30 | 63.13 | 60.24 | 0.50 | 90.8 | 89.45 | 87.34 | 87.38 | 0.52 |
| | (21.16) | (24.19) | | (28.2) | (27.07) | (28.04) | | (14.79) | (15.58) | (18.70) | (16.99) | |
| Loss in Last Year (%) (self-report) | 6.72 | 6.98 | 0.92 | 24.29 | 22.96 | 23.01 | 0.79 | 31.60 | 29.38 | 26.94 | 29.37 | 0.53 |
| | (15.14) | (16.91) | | (15.41) | (15.12) | (15.33) | | (18.02) | (15.30) | (13.65) | (17.51) | |
| **Panel B:after playing the game** | | | | | | | | | | | | |
| Risk Aversion | | | | 4.13 | 4.16 | 4.10 | 0.95 | 3.20 | 3.23 | 3.04 | 3.11 | 0.90 |
| | | | | (1.45) | (1.44) | (1.43) | | (1.52) | (1.44) | (1.59) | (1.71) | |
| Perceived Probability of Future Disaster | | | | 23.10 | 22.33 | 21.64 | 0.76 | 24.10 | 23.15 | 21.38 | 23.80 | 0.30 |
| | | | | (15.77) | (15.52) | (14.53) | | (9.83) | (9.26) | (9.26) | (9.38) | |
| Take-up(%) | 0.19 | 0.24 | 0.42 | 0.17 | 0.17 | 0.32 | 0.01 | 0.28 | 0.39 | 0.37 | 0.36 | 0.61 |
| | (0.39) | (0.43) | | (0.38) | (0.38) | (0.47) | | (0.45) | (0.49) | (0.49) | (0.48) | |
| Observations | 86 | 95 | | 121 | 124 | 134 | | 52 | 73 | 49 | 151 | |

Note: standard deviations are in the parentheses.

[38]We did not ask these questions before the games; if players had decided to act consistently with their answers, this would have obscured the treatment effects.

[39]The enumerators gave sample individuals 10 small paper balls and asked them to put these paper balls into two areas: (1) no disaster reducing yield more than 30% next year and (2) disaster reducing yield more than 30% next year. If households put two paper balls into area (2) and eight paper balls into area (1), their perceived probability of future disaster is around 20%.

*P-value in wave 1 is for F test of equal means of two groups

** P-values in wave 2 and 3 are for Wald test of equal means of three and four groups

***Education is coded as follows: 0-illitracy; 1-primary school; 2-secondary school; 3-high school; 4-College

Table 1 presents summary statistics and balance checks separately for each wave. In total, we visited 885 households in round 1 and 816households in round 2. The overall attrition rate between round 1 and round 2was 7.8 percent. While the attrition was slightly higher in the game group, 9.8 percent, than in the control and calculations groups, respectively 6.2 and 5.6 percent, the difference in attrition between groups is not statistically significant. Attrition was 11.8 percent in the information group, which is not significantly different from the 10.4 percent attrition in the no information group in wave 3.

The summary statistics show that household heads are almost exclusively male. The average education level is between primary school and secondary school. The average individual is risk averse. The randomization check shows that most control variables are balanced. The only exception is that in wave 1, the households in the game group are older than those in the control group. However, the regressions in the next section show that the relationship between take-up and age is in any case insignificant.

## 4. Empirical Result

### 4.1 The Impact of Hypothetical Experience on Actual Take-up

In what follows, "Game" refers to households who were assigned to the *Game 20%* group or the *Game 10%* group. As shown in Figure 4, the take-up rate of the control group is 19.8 percent, while that of the calculation group is 24.7 percent. In the game group, the take-up is 32.3 percent. Thus, both the game and the calculation treatment increase take-up, but the game treatment is more effective.

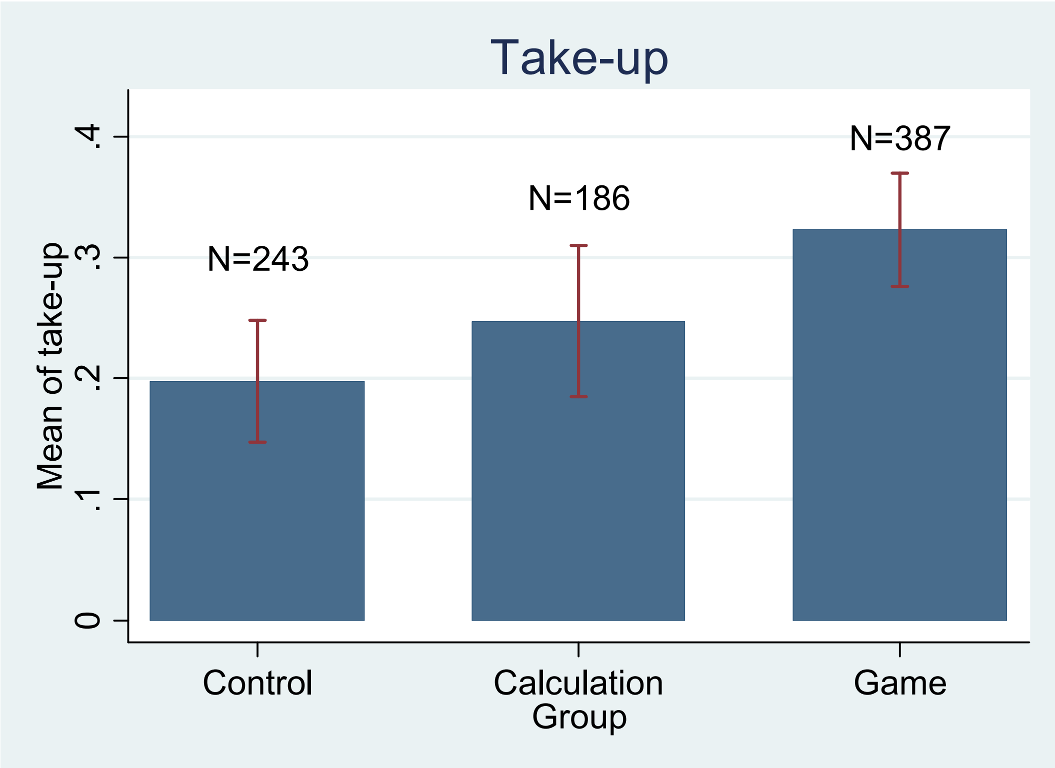



Figure 4 Treatment effect

Note: This figure shows the treatment effect of the calculation group and the game group. In the control group, the take-up is 19.8%. In the calculation group, the take-up increases to 24.7%.In the game group, the take-up increases to 32.3%. It suggests that both the game treatment and the calculation treatment increase the actual take-up and the game treatment is more effective.

Figure 5 shows the treatment effects of the game treatment and the calculation treatment when interacted with the information treatment. In the no information group, the pattern is similar to Figure 4. However, the game treatment increases the take-up and is more effective than the calculation treatment. In the information group, the take-up rates of three groups are similar. This suggests that the game treatment and the calculation treatment are not as effective with the interaction of information treatment.

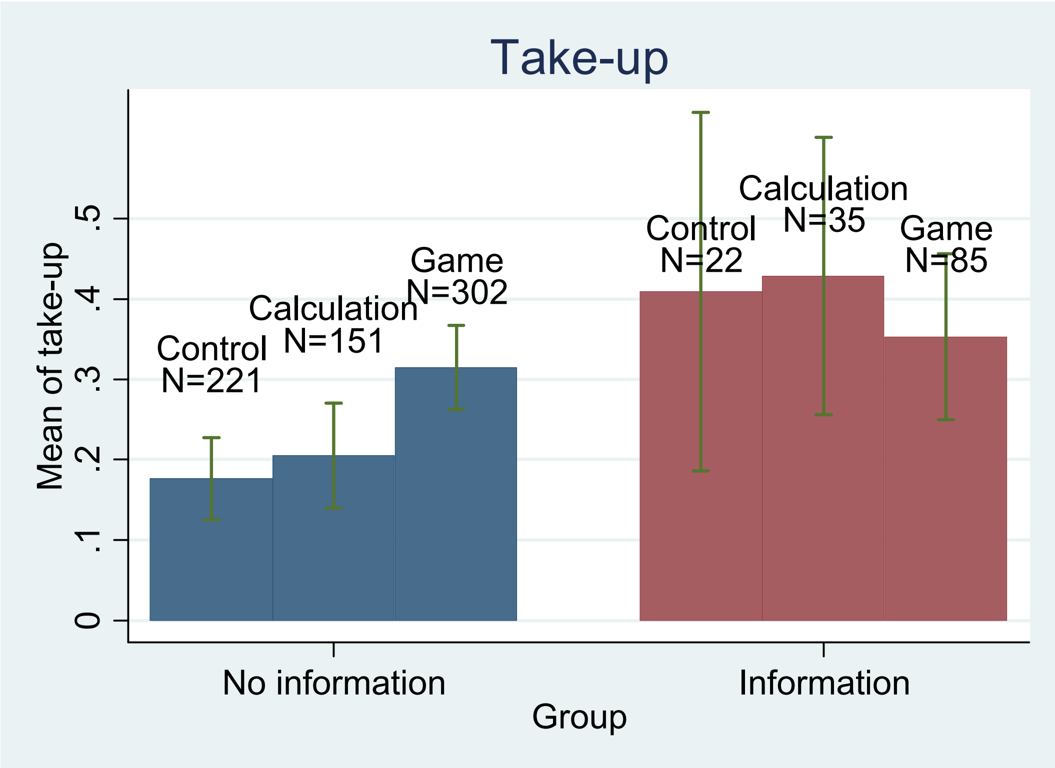



Figure 5 Treatment effect by the information treatment

Note: This figure shows the treatment effect by the information treatment. Without the information treatment, the game treatment is more effective than the calculation treatment. With the information treatment, the game treatment and the calculation treatment is not effective.

We estimate the treatment effect on the take-up decision through a logit regression in (1):

$$buy_{ij} = \alpha_j + \alpha_k + \beta_g Tg_{ij} + \beta_c Tc_{ij} + \phi X_{ij} + \varepsilon_{ij} \quad (1)$$

where $buy_{ij}$ is an indicator that takes on a value of one if household $i$ in natural village $j$ buys the insurance. $Tg_{ij}$ is an indicator for the game treatment and $Tc_{ij}$ is an indicator for the calculation treatment. Random assignment implies that $\beta_g$ is an unbiased estimate of the reduced-form intention-to-treat (ITT) game treatment effect and $\beta_c$ is an unbiased estimate of the ITT calculation treatment effect. $X_{ij}$ are household characteristics (e.g. , gender, age, years of education, household size, land for production, whether they own a car, etc) and $\alpha_j$ and $\alpha_k$ are village fixed effects and enumerator fixed effects, respectively. $\varepsilon_{ij}$ is type I extreme value error term. Since our roll-out design has three waves, it is important to control for potential confounding variables such as the covariates (X) and fixed effects. We report marginal effects in Table 2.

**Table 2. The Effect of Game and Calculation on Insurance Take-up**

| **Specification:** | **Logistic regression** | | | | |
|---|---|---|---|---|---|
| **Dep. Var.:** | **Individual adoption of insurance** | | | | |
| **Sample:** | **All Sample** | **No Information** | **Information** | **All Sample** | |
| | 1 | 2 | 3 | 4 | 5 |
| Game | 0.092 | 0.119 | -0.086 | 0.096 | 0.092 |
| | (0.039)** | (0.034)*** | (0.172) | (0.037)*** | (0.038)** |
| Calculation | 0.025 | 0.012 | -0.009 | 0.029 | 0.031 |
| | (0.043) | (0.047) | (0.189) | (0.042) | (0.040) |
| %Loss Last Year (self report) | | | | 0.207 | 0.200 |
| | | | | (0.104)** | (0.110)* |
| Age | | | | | 0.008 |
| | | | | | (0.011) |
| Education | | | | | 0.039 |
| | | | | | (0.017)** |
| Household Size | | | | | -0.015 |
| | | | | | (0.005)*** |
| Land of Rice Production | | | | | 0.002 |
| | | | | | (0.014) |
| Wald Test: $\beta_g=\beta_c$ | | | | | |
| p-value | 0.1376 | 0.0117** | 0.5376 | 0.1328 | 0.1568 |
| Obs. | 816 | 674 | 132 | 816 | 816 |
| Omitted Treatment | Control | | | | |
| Mean of Dep. Var. for Omitted Treatment: | 0.198 | | | | |
| Fixed Effects for Village and Enumerator | Y | Y | Y | Y | Y |
| Log Likelihood | -431 | -335 | -86 | -429 | -424 |
| Pseudo R-square | 0.0918 | 0.1057 | 0.0323 | 0.0962 | 0.1065 |

Notes: Dependent variable is individual adoption; standard errors are clustered by 16 natural villages. Robust clustered standard errors are in the parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level. In column 2, we restrict the sample to households in the no information group. In column 3, we restrict the sample to households with the information treatment. In column 4 to 5, we add dummies for missing values of control variables in the regression. In column 4, the self reported percentage of loss in last year is included in the regression. In column 5, additional control variables are age group of household head, education of household head, household size and area of rice production. We lose ten observations in column 3 because one independent variable predicts not buying perfectly and the logistic regression drops them.

Column 1 presents results from the simplest possible specification, where the only right hand side variables are the indicators for the game treatment, the calculation treatment, and the village and enumerators fixed effects. The marginal effect of the game treatment (0.096) is positive and significant at the 5% level. Thus, the game treatment increases the take-up by 9.6 percentage points, which is about a 48 percent increase relative to the baseline take-up of 20 percentage points. The marginal effect of the calculation treatment (0.027) is positive but it is not statistically significant.

In column 2, we restrict the sample to households in the no information group. The marginal effect of the game treatment (0.126) increases and the pattern is similar to column 1.

In column 3, we restrict the sample to households in the information group. The marginal effects of the game and calculation treatment are imprecisely estimated; they are negative and not statistically significant. The difference in marginal effects between the information group and the no information group is significant at the 10% level.

In column 4, the self reported percentage of loss last year and a dummy for missing values are included in the regression with all samples. The pattern is similar to column 1.The marginal effect of the percentage of loss last year is 0.22%; this is significant at the 10% level. Thus, the effect of playing games is roughly of the same magnitude as the effect of a 45 percentage point increase in actual loss last year.

In column 5, a variety of other control variables and dummies for missing values are additionally included in the regression with all samples. The pattern is still similar to column 1.Education level is positively correlated with take-up and household size is negatively correlated with the take-up.

In sum, the game treatment increases the insurance take-up by 9 to 10 percentage points, resulting in an increase of around 45 to 50 percent relative to baseline take-up of 20 percentage points. The effect of playing games is roughly of the same magnitude as a 45 percentage point increase in actual loss during the previous year.

### 4.2 Possible Channels

In order for these findings to illustrate channels, more information is needed to analyze the mechanisms through which this effect could work. Possible explanations include: (1) changes in risk attitudes, (2) changes in the perceived probability of future disasters, (3) learning about the benefits of insurance, or (4) changes in hypothetical experience of disasters.

#### 4.2.1 Risk Attitudes

First, it is possible that the treatment increases take-up by changing risk attitudes. To determine whether the game treatment changes risk attitudes and increases take-up, we run the follow regressions to test it:

$$buy_{ij} = \alpha_{2j} + \beta_{risk}\, risk_{ij} + \beta_{prob}\, prob_{ij} + \delta_{ij} \quad (2)$$

$$risk_{ij} = \alpha_{3j} + \gamma_{gr} Tg_{ij} + \gamma_{cr} Tc_{ij} + \eta_{ij} \quad (3)$$

$$risk_{ij} = \alpha_{4j} + \beta_{dr}\, disaster_{ij} + \omega_{ij} \quad (4)$$

where $risk_{ij}$ is an increasing measurement of risk aversion and $disaster_{ij}$ is the number of hypothetical disasters households experienced during the games. Equation (2) analyzes the correlation between take-up and risk attitudes. We restrict the sample to the control group and the calculation group in Equation (2) because we asked them questions about their risk attitudes and the perceived probability of future disasters before any intervention took place. In Equation (3) and (4), we estimate the effects of playing games and experiencing disasters, respectively. We assume that there is no measurement error as to risk attitudes and perceived probability of future disaster, and that the estimation in Equation (2) is unbiased.

**Table 3. The Decomposition Effect of Game and Calculation**

| **Specification:** | **OLS Regression** | | | | |
|---|---|---|---|---|---|
| **Dep. Var.:** | **Individual Adoption of Insurance** | **Risk Aversion** | | **Perceived Probability of Future Disaster** | |
| **Sample:** | **Control & Calculation** | **All Sample** | **Game** | **All Sample** | **Game** |
| | 1 | 2 | 3 | 4 | 5 |
| Risk Aversion | 0.035 | -0.024 | | | |
| | (0.016)** | (0.182) | | | |
| Perceived Probability of Future Disaster | 0.215 | 0.055 | | | |
| | (0.110)* | (0.165) | | | |
| Game | | | | -0.015 | |
| | | | | (0.008)* | |
| Calulation | | | | -0.011 | |
| | | | | (0.009) | |
| Number of Hypothetical Disasters | | | 0.080 | | 0.003 |
| | | | (0.138) | | (0.008) |
| Obs. | 329 | 697 | 320 | 667 | 310 |
| Omitted Treatment | | | Control | | |
| Mean of Dep. Var. for Omitted Treatment: | 0.198 | Y | Y | Y | Y |
| Social-economic Variables | Y | Y | Y | Y | Y |
| Fixed Effects for Village and Enumerator | Y | Y | Y | Y | Y |
| R-square | 0.1397 | 0.1932 | 0.2022 | 0.0990 | 0.1896 |

Notes: Standard errors are clustered by 16 natural villages. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. In column 1, we restrict the sample to the control group and the calculation group and regress adoption on risk attitude. In column 2 to 3, we regress risk attitude on treatment indicator and controls. In column 4 to 5, we regress the perceived probability of future disasters on treatment indicator and controls.

In column 1 of Table 3, estimates from (2) are presented. The coefficient of risk aversion (0.032) is positive and significant at the 5% level. The coefficient of perceived probability of future disasters (0.0214) is positive and significant at the 10% level. Column 2 presents the estimates of (3), including various controls and dummies for missing values. Column 3 restricts the sample to households who played the hypothetical games and presents the estimates of (4). The results show that the treatment has no effect on risk aversion and the coefficient of the number of hypothetical disasters is not statistically significant.

To determine whether the game treatment changes risk attitudes and increases take-up, we stack Equation (1), (2), and (3), generate indicators for each equation, and estimate the regression system. To account for the correlation of error terms between each equation, standard errors are clustered by 16 natural villages. We test the hypothesis: $\beta_{risk}\gamma_{gr} = \beta_g$. We reject the hypothesis at the 5% level (p=0.039), with the 95% confidence interval ranging in [-0.013, 0.011]. To determine whether the number of hypothetical disasters changes risk attitudes and increases take-up, we stack Equation (1), (2), and (4) and estimate the regression system. We test the hypothesis: $1.48\beta_{dr}\gamma_{gr} = \beta_g$, where 1.48 is average number of hypothetical disasters experienced during the games. We reject the hypothesis at the 5% level (p=0.044), with the 95% confidence interval of $1.48\beta_{dr}\gamma_{gr}$ ranging in [-0.004, 0.004]. These results suggest that changes in our measurement of risk attitudes are unlikely to explain our main treatment effect.

**4.2.2 The Perceived Probability of Future Disaster**

Demand for insurance also depends on the perceived probability of future disasters. It is possible that the games increase take-up by changing the perceived probability of future disasters. To test this channel, we run the following regressions:

$$prob_{ij} = \alpha_{3j} + \gamma_{gp} Tg_{ij} + \gamma_{cp} Tc_{ij} + \eta_{ij} \qquad (5)$$

$$prob_{ij} = \alpha_{4j} + \beta_{dp}\, disaster_{ij} + \omega_{ij} \qquad (6)$$

where $prob_{ij}$ is the perceived probability of future disaster. In Equation (5) and (6), we estimate the effects of playing games and experiencing disasters, respectively. The results of (5) and (6) are presented in column 4 and 5 in Table 3, respectively.

The treatment has a negative effect on the perceived probability of future disasters in columns 4. The coefficient of the number of hypothetical disasters is not significant. Following a similar procedure as in section 4.2.1, we test the hypothesis $\beta_{prob}\gamma_{gp} = \beta_g$ and $1.48\beta_{dp}\gamma_{gp} = \beta_g$ to determine whether changes in the perceived probability of future disasters is the channel. We reject that at the 5% level.

To determine whether the total effects of changing risk attitudes and the perceived probability of future disasters are the channel through which the observed effects operate, we follow a similar procedure as in section 4.2.1 and test the following two hypothesis: $\beta_{risk}\gamma_{gr} + \beta_{prob}\gamma_{gp} = \beta_g$ and $1.48\beta_{dr}\gamma_{gr} + 1.48\beta_{dp}\gamma_{gp} = \beta_g$. We reject the hypothesis at the 5% level. These results suggest that the total effects of changes in risk attitudes and the perceived probability of future disasters are unlikely to explain our main treatment effect.

### 4.2.3 Learning the Benefits of Insurance

It is also possible that playing insurance games provides direct information about the benefits of insurance. To test that, we compare the treatment effect of the game and calculation treatment; the difference between those two interventions should indicate whether households acquire disaster experiences during the games.

We run various regressions with (1) and report the p-value of Ward test $\beta_g = \beta_c$ in Table 2. In columns 1, 4 and 5, we use the whole sample. The difference between $\beta_g$ and $\beta_c$ is around 7 percentage points and it is not statistically significant (p-value of Ward test is between 0.13 and 0.16). In columns 2, we restrict the sample to the no information group. The difference between $\beta_g$ and $\beta_c$ is around 11 percentage points and is significant at the 5% level.

In sum, when we consider the channel of the game treatment effect without the interaction effect of the information treatment, we conclude that learning about the benefits of insurance is unlikely to explain the treatment effect of playing games. When we consider the channel of the game treatment effect and interaction effect of the information treatment, there is suggestive evidence that learning about the benefit is unlikely to explain the game treatment effect.

### 4.2.4 The Experience of Hypothetical Disasters

Another hypothesis is that hypothetical experience matters. To test this hypothesis, we explore the randomization of the number of hypothetical disasters during the game. We present Figure 6 about actual take-up and the hypothetical disasters experienced during the games.

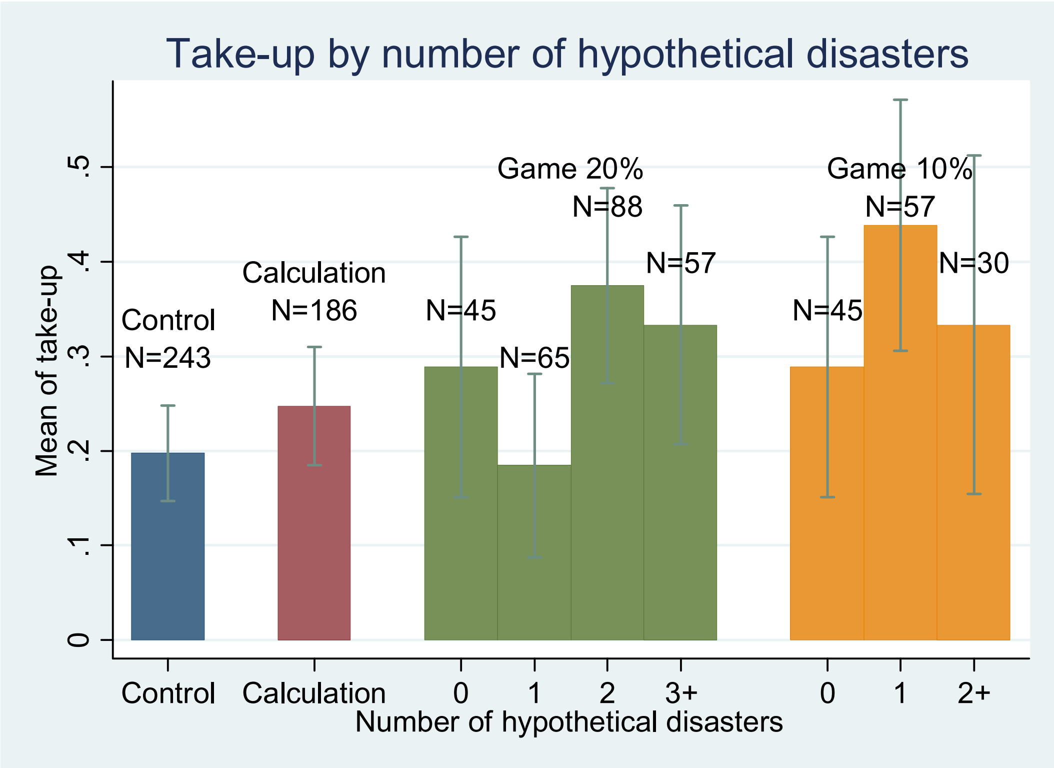



Figure 6: Take-up by number of hypothetical disasters in the games

Note: the figure shows the insurance take-up conditioning on the number of disasters they experienced during the games. The left two bars show the take-up of the control group and the calculation group.

In the *Game 20%* group, the take-up among households who experienced two or more disasters is higher than that among those who experienced zero or one disaster. In the *Game 10%* group, the take-up of households who experienced one disaster is higher than those who experienced either zero or two and more disasters. However, given the relatively large standard deviation, Figure 6 provides only suggestive evidence that the take-up rate is increasing in the number of hypothetical disasters experienced and that the take-up in the group with no hypothetical disasters is greater than that in the control group.

To understand this further, we run the following regression:

$$buy_{ij} = \alpha_j + \beta_{disaster} \; disaster_{ij} + \delta_{ij} \qquad (7)$$

where $disaster_{ij}$ is the number of hypothetical disasters experienced during the games.

**Table 4. the Effect of Hypothetical Games on Actual Insurance Take-up**

| **Specification:** | **Logistic Regression** | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | **Individual Adoption of Insurance** | | | | | |
| **Sample:** | **All sample** | | | **No information** | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Game | 0.010 | | 0.047 | 0.037 | | 0.085 |
| | (0.059) | | (0.046) | (0.067) | | (0.047) |
| Calculation | 0.042 | | 0.044 | 0.032 | | 0.037 |
| | (0.046) | | (0.045) | (0.051) | | (0.050) |
| Number of Hypothetical Disasters | 0.059 | | | 0.055 | | |
| | (0.031)* | | | (0.036) | | |
| Game and No Disaster | | 0.030 | | | 0.060 | |
| | | (0.060) | | | (0.076) | |
| Game and One Disaster | | 0.046 | | | 0.064 | |
| | | (0.045) | | | (0.044) | |
| Game and Two Disasters | | 0.137 | | | 0.159 | |
| | | (0.043)*** | | | (0.042)*** | |
| Game and Three or More Disasters | | 0.133 | | | 0.143 | |
| | | (0.066)** | | | (0.062)** | |
| Number of Hypothetical Disasters in First Half of Game (2011-2015) | | | -0.019 | | | -0.042 |
| | | | (0.024) | | | (0.028) |
| Number of Hypothetical Disasters in Second Half of Game (2016-2020) | | | 0.070 | | | 0.072 |
| | | | (0.033)** | | | (0.034)** |
| Obs. | 804 | 804 | 804 | 664 | 664 | 664 |
| Omitted Treatment | Control | | | | | |
| Mean of Dep. Var. for Omitted Treatment: | 0.198 | | | | | |
| Social-economic Variables | Y | Y | Y | Y | Y | Y |
| Fixed effects for village and enumerator | Y | Y | Y | Y | Y | Y |
| Log Likelihood | -427 | -427 | -426 | -333 | -334 | -331 |
| Pseudo R-square | 0.0599 | 0.0864 | 0.0884 | 0.0956 | 0.0965 | 0.1021 |

Notes: Dependent variable is individual adoption; standard errors are clustered by 16 natural villages. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. In column 4 to 6, we restrict the sample to households in the no information treatment. In column 3 and 6, we regress the actual take-up on the number of hypothetical disasters in the first 5 rounds and the number of hypothetical disasters in the last 5 rounds.

The marginal effect estimated in (7) is presented in column 1 and 4 of Table 4. In column 1, the coefficient (0.059) is positive and statistically significant at the 10% level. In the no information group (column 4), the coefficient (0.055) is positive but not significant (p=0.127). Therefore, the results suggest that the treatment effects were driven mainly by those who experienced more hypothetical disasters during the games.

Hypothetical experience might change two things: understanding about insurance and vividness. We run regression in Equation (8) to analyze these two effects:

$$buy_{ij} = \alpha_j + \beta_0 disaster\ 0_{ij} + \beta_1 disaster\ 1_{ij} + \beta_2 disaster\ 2_{ij} + \beta_3 disaster\ 3_{ij} + \varepsilon_{ij} \quad (8)$$

where $disasterK_{ij}$ is an indicator that takes on a value of one if households experience $K$ disasters during the games. $\beta_0$ captures the understanding effect; the difference between $\beta_0$ and other coefficients captures the vividness effect.

The marginal effect of (8) is presented in column 2 and 5 in Table 4. The coefficients of $disaster0_{ij}$ and $disaster1_{ij}$ are positive but not statistically significant. Indicators for more disasters are positive, statistically significant and relatively large in magnitude. In the no information group (column 4), the coefficients are relatively larger, which is similar to what we have seen in Table 2. The difference between $\beta_1$ and $\beta_2$ is statistically significant at the 10% level. However, we cannot reject the hypothesis that $\beta_0$ and $\beta_1$ are the same. Therefore, we cannot distinguish between the understanding effect and the vividness effect.

To further understand how hypothetical experience influences take-up, we present the take-up conditioning on disaster in the first 5 rounds and in the last 5 rounds in Figure 7.

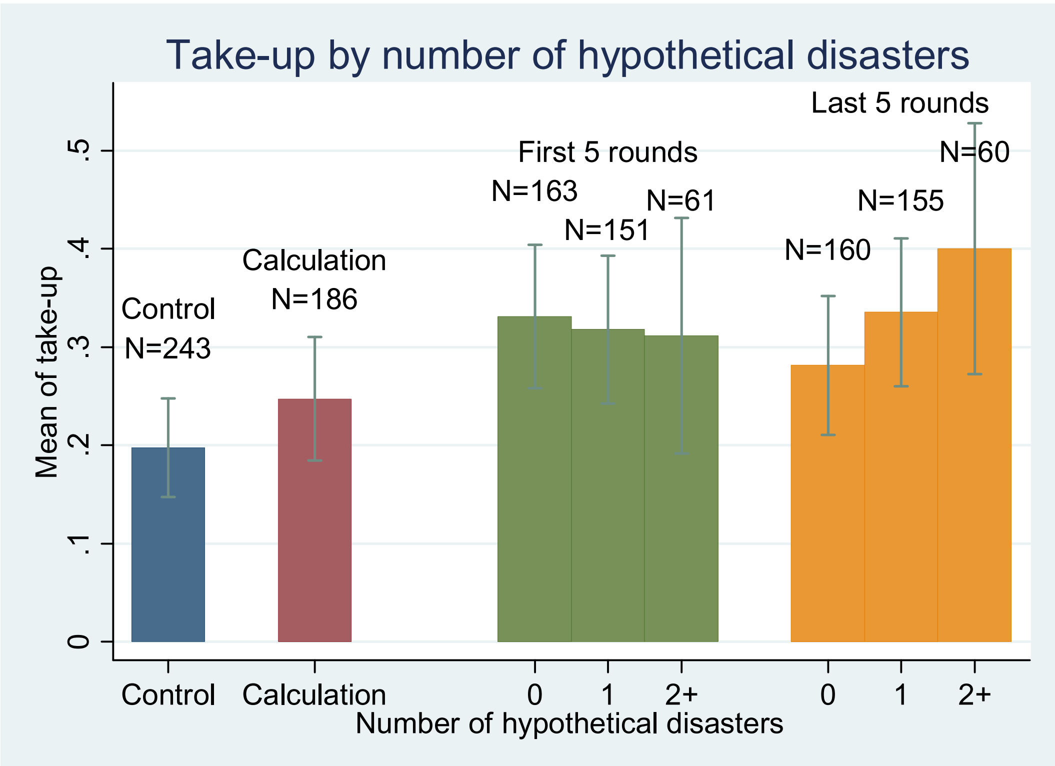



Figure 7: Take-up by number of hypothetical disasters in different rounds

Note: the left two bars show that the insurance take-up conditioning whether there is a disaster in the first round and last round. The right two bars show the insurance take-up conditioning on the number of disasters in the first 5 rounds and last 5 rounds.

The evidence in Figure 7 suggests that the number of hypothetical disasters experienced in the first 5 rounds does not influence take-up, whereas the number of disasters in the last 5 rounds appears to have a bigger effect.

We then create two new variables: the number of hypothetical disasters in the first 5 rounds and the number of hypothetical disasters in the last 5 rounds. We run the following regression:

$$buy_{ij} = \alpha_j + \beta_{f5} disaster\_first5_{ij} + \beta_{l5} disaster\_last5_{ij} + \delta_{ij} \quad (9)$$

As seen in column 4, the coefficient of "disasters in the first half" is negative and not statistically significant. However, the coefficient of "disasters in the last half" is positive and significant at the 5% level. The coefficient suggests that experiencing an additional disaster in the last half increases take-up by 7.0 percentage points. In the no information group (column 6), the coefficient of the last 5 rounds is also positive and significant at the 5% level. This pattern is robust to different measurement of the first and last few rounds. If we regress take-up on the number of hypothetical disasters in the first (10-n) rounds and that in the last n rounds, we find that when n equals 5,6,7,8 or 9, the coefficients of the last n rounds are positive and significant at the 5% level.[40] These results are consistent with the literature in experienced utility and recency effects (Fredrickson and Kahneman 1993; Schreiber and Kahneman, D. 2000), where they find that the affect experienced during the last moments of the experiment has a privileged role in subsequent evaluations, and late moments in the experiment are assigned greater weight than earlier ones.

To summarize, we find that both the total number of disasters and the number of disasters in last few rounds increase take-up significantly. These results suggest that the experience of recent hypothetical disasters might be the channel through which the games influence insurance decisions.

## 5. Models

The evidence so far implies that hypothetical experience influences the actual insurance decisions. In this section, we present a simple model to illustrate how such an effect could occur. In section 5.1, we show that standard constant absolute risk aversion (CARA) preferences and constant relative risk aversion (CRRA) preferences are unlikely to explain the data. In section 5.2, we add a weight parameter to the utility function to capture the influence of experience. Then we estimate the parameters through a maximum likelihood method (MLE).

### 5.1 Standard Model

We first consider a simple model with CARA preferences commonly used in the insurance literature (Einav *et al.* 2010).

$$u(x) = -\frac{\exp(-\alpha x)}{\alpha} \quad (11)$$

With CARA preferences, the consumer's wealth does not affect his insurance choices. Therefore, the take-up decisions should be determined by the joint

[40] See Appendix Table A4 for detail

distribution of risk attitudes and perceived probability of future disasters.

Let $U(a)$ denote the household utility as a function of the insurance decision. $a=1$ if the household buys the insurance and $a=0$ if the household does not buy the insurance. Let $(b,\tau)$ denote the insurance contract in which $b$ is the repayment of insurance if there is a disaster and $\tau$ is the premium. Let $x$ be the gross income of rice production and $p$ the perceived probability of future disasters. Let $l$ denote the loss in yield. The expected utility of not buying the insurance is:

$$U(a=0)=(1-\mathrm{p})u(x)+\mathrm{p}u(x-l) \quad (12)$$

If a household buys insurance, it should earn its normal income and pay the premium when there is no disaster; it should have a loss and receive payment from the insurance company when there is a disaster. The utility of buying the insurance is:

$$U(a=1)=(1-\mathrm{p})u(x-\tau)+\mathrm{p}u(x-\tau-l+b) \quad (13)$$

The condition for the household to buy the insurance is

$$U(a=1)>=U(a=0) \quad (14)$$

It is straightforward to show that the households who are more risk averse and whose perceived probabilities of future disasters are larger are more likely to buy the insurance.

To test whether the standard CARA preferences could explain our data, one way is to use the parameter as measured, calibrate individual decisions and compare the calibrated decisions with actual decisions. We assume that there is no measurement error for risk aversion ($\alpha$) or for the perceived probability of future disasters (p). Although we do not observe parameter $\alpha$, we can make use of the choices in Table 1 to estimate the intervals of their $\alpha$ in the utility function. The intervals of $\alpha$ under CARA and CRRA are presented in Table 5. If a household takes two riskless options, $\alpha$ should be greater than zero and less than 0.0041 under CARA preferences. The details of the simulation procedures are discussed in Appendix C.

**Table 5. Range of Risk Aversion Parameter**

| Number of Riskless Options Taken | Range α of for CARA $u(x)=-\exp(-\alpha x)/\alpha$ | Range α of for CRRA $u(x)=x^{1-\alpha}/(1-\alpha)$ |
|---|---|---|
| 0 | α<-0.0121 | α<-1.4 |
| 1 | -0.0121<α<-0.0041 | -1.4<α<-0.35 |
| 2 | -0.0041<α<0 | -0.35<α<0 |
| 3 | 0<α<0.0041 | 0<α<0.25 |
| 4 | 0.0041<α<0.0121 | 0.25<α<0.5 |
| 5 | α>0.0121 | α>0.5 |

Notes: Calculation of range of risk aversion parameter is based on the number of riskless options taken in table A1.

We find that the mean of simulated take-up is 81.08% and the standard deviation is 0.0049. This contradicts our actual data that the take-up in the sample is 26.84%. This suggests that standard CARA and CRRA preferences are unlikely to explain our data.

Another route is to ignore $\alpha$ and $p$ as elicited. Suppose that we had not elicited measures for risk aversion and perceived probability of future disasters. Then we estimate $\alpha$ and p in the logit formula (15) through MLE:

$$P(a=1)=\frac{\exp(U(a=1))}{\exp(U(a=1))+\exp(U(a=0))} \qquad (15)$$

We find that the model is not identifiable. The log-likelihood function reaches a flat region and the combination of $\alpha$ and $p$ falls into the following two categories: (1) negative $\alpha$ (risk seeking) and $p$ greater 17% (2) positive $\alpha$ (risk averse) and $p$ less than 5%. This contradicts our data that average risk attitude implies risk aversion and that the average perceived probability of future disasters is around 20%.

In sum, both the calibrated decisions and the estimated parameters contradict our data under standard CARA and CRRA preferences. These results suggest that standard CARA and CRRA preferences are unlikely to explain the observed take-up rates in the presence of the perceived probability of future disasters which our questions elicited.

**5.2 Model Based on Experience**

We have shown that standard CARA and CRRA preferences are unlikely to explain the data. In order to develop a model that fits our data, we add a weight parameter to capture the effect of experience. It is possible that households buy more insurance because they pay more attention to disasters and benefits after they experience the hypothetical disasters during the games. We develop a simple model in the following.

$$U(a=0)=(1\text{ - p})u(x)+\text{p}u(x-\mu l) \qquad (16)$$

$$U(a=1)=(1\text{ - p})u(x-\tau)+\text{p}u(x-\tau-\mu l+\mu b) \qquad (17)$$

where $\mu$ is a parameter that measures the weight of disaster loss and insurance benefits. The idea is that households might give less weight to disasters and benefits which they experience infrequently. When they are treated with games, they experience disasters and insurance benefits during the hypothetical games. These hypothetical disasters draw their attention to disaster loss and insurance benefits and increase the weight parameter $\mu$.

It is straightforward to show that, under the assumption of CARA preferences with inattention parameter $\mu$, if $\alpha > 0$, then $\frac{\partial P(buy==1)}{\partial \mu} > 0$. To the extent that playing games increases $\mu$, it would increase the insurance take-up. To test this, we allow $\mu$ in the group who do not play games ($\mu_1$) to be different from $\mu$ in the group who play games ($\mu_2$). Then we estimate $\mu_1$ and $\mu_2$ with MLE and simulation. The details of the estimation procedures are discussed in Appendix C.

**Table 6. Maximum Likelihood Estimation of Utility Function**

| | CARA | | | CRRA | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| $\mu_1$ | 0.208 | 0.204 | | 0.152 | 0.149 | |
| $\mu_2$ | 0.370 | 0.339 | | 0.269 | 0.262 | |
| δ | | -1.097 | | | -0.689 | |
| c | | | 0.203 | | | 0.205 |
| $k_a$ | | | 0.075 | | | 0.012 |
| $k_h$ | | | 4.254 | | | 0.735 |
| 90% CI for $\mu_1$ or $k_a$ | [0.106,0.391] | [0.121,0.395] | [0.000,0.450] | [0.121,0.203] | [0.121,0.174] | [0.000,0.082] |
| 90% CI for $\mu_2$ or $k_h$ | [0.152,0.645] | [0.152,0.546] | [0.000,32.689] | [0.152,0.645] | [0.174,0.311] | [0.000,2.326] |
| t test p-value | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** |
| k-s test p-value | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** |
| Obs. | 613 | 613 | 344 | 613 | 613 | 344 |
| Number of Draws for α | 100 | 100 | 100 | 100 | 100 | 100 |

Notes: we estimate parameters in CARA utility function $u(x) = -\frac{\exp(-\alpha x)}{\alpha}$ and CRRA utility function $u(x) = \frac{x^{1-\alpha}}{1-\alpha}$ through MLE. In all columns, we constrain $\alpha$ to be uniform draws from the intervals of their risk attitudes and constrain $p$ to be the perceived probability of future disasters from our survey data. We present the mean of coefficients from 100 draws of $\alpha$. In column 1 and 4, we allow the weight parameter in the group who do not play games ($\mu_1$) to be different from weight parameter in the group who play games ($\mu_2$). In column 2 and 5, we add $\delta$ to measure the utility of understanding the insurance if they buy the insurance. We normalize $\delta$ to be zero in the game treatment so that the estimated $\delta$ is the difference of the utility of understanding. In column 3 and 6, we assume that the weight parameter has the following structure $\mu = C + D(1-\exp(-k_a f_a - k_h f_h))$. Then we estimate both the learning effect of actual experience ($k_a$) and hypothetical experience ($k_h$) with different measurement of actual disaster.

The result is presented in column 2 table 6. We find that the estimated mean of $\mu_1$ is around 0.21 and that $\mu_2$ is around 0.37. The T-test and Kolmogorov-Smirnov test show that the mean and the distribution are significantly different at the 1% level.

Column 6 presents the result with CRRA preferences. Although the point estimates are different, the key pattern is similar. These results are consistent with our hypothesis that playing games increases $\mu$ and thus increases insurance take-up.

Hypothetical experience might have two effects: changes in understanding and changes in vividness. We add another parameter $\delta$ in (17):

$$U(a = 1) = (1 - p)u(x - \tau) + pu(x - \tau - \mu l + \mu b) + \delta \quad (18)$$

where $\delta$ measures the utility of understanding insurance if they buy the insurance. The intuition is that households would be less happy if they buy something they do not understand than something they understand. It might capture ambiguity aversion and it is a reduced form in our model. We normalize $\delta$ to be zero in the game treatment so that the estimated $\delta$ is the difference of the utility of understanding. We estimate $\mu_1$, $\mu_2$ and $\delta$ with the same procedure to estimate $\mu_1$ and $\mu_2$. The results are presented in column 3.

The estimated mean of $\mu_1$ is about 20.4% and $\mu_2$ is about 33.9%. The T-test and Kolmogorov-Smirnov test show that the mean and the distribution are significantly different at the 1% level. The estimated mean of $\delta$ is -1.097 and the t-test shows that the mean is significantly different from zero at the 1% level. Since we normalize $\delta$ to be zero in the game treatment, this means that the utility of understanding is higher in the game treatment. Column 7 presents the result with CRRA preference. Although the point estimates are different, the key pattern is similar. These results are consistent with our hypothesis that playing games increases the understanding and vividness and thus increases the insurance take-up.

In order to understand the empirical relationship between experience and the weight parameter, we model $\mu$ following the lead of Agarwal *et al.* (2011).

$$\mu = C + D(1 - \exp(-k_a f_a - k_h f_h))$$

where $k_a, k_h, C, D > 0$, and $C + D \le 1$.

$f_a$ is actual experience, measured by percentage of disasters reducing yield more than 30% in the past 3, 2 or 1 years. $f_h$ is hypothetical experience, measured by percentage of disasters during 10 rounds of games. $k_a$ and $k_h$ capture the rate of learning from actual experience and hypothetical experience. With enough experiences, attention saturates to C + D. If $C + D = 1$, attention is perfect in the long run, but if $C + D < 1$, attention is imperfect, even in the long run. Here, we assume $C + D = 1$. Then we could estimate $k_a$ and $k_h$ and compare the effect of actual and hypothetical experience.

In column 4, we estimate the learning effect of both actual experience ($k_a$) and hypothetical experience ($k_h$) under CARA preferences. $f_a$ is measured by percentage of disasters reducing yield more than 30% in the past 3 years. The mean of $k_a$ is 0.075 and the mean of $k_h$ is 4.254; they are significantly different at the 1% level. Column 8 presents the result with CRRA preferences. Although the point estimates are different, the key pattern is similar. These results suggest that both actual and hypothetical experience matter. Moreover, experience acquired in the recent insurance game has a stronger effect on the actual insurance take-up than that of real disasters in the previous year.

**6. Conclusion**

It is important to understand why the take-up for weather insurance is low even when farmers face substantial natural risks. We apply a novel method of financial education and test for the role of experience and information in weather insurance take-up in rural China. We find that playing insurance games increases the actual insurance take-up by 9.6 percentage points, a 48% increase relative to the baseline take-up of 20 percentage points. We investigate the possible mechanisms through which this effect could work, and find that changes in experience of disasters and insurance benefits are very likely to be the channel.

There is mixed evidence in the literature as to whether financial education is effective to change individual decisions. Why is financial education effective sometimes but not others? Under what circumstances is financial education effective? This paper shows that financial education with simulated experiences can help increase insurance take-up in rural areas. Gaurav et al. (2011) finds similar results in India. Song (2011) finds that learning the concept of compound interest has a positive and significant effect on weather insurance adoption in rural China. These suggest that we should first identify the barriers to individual participation and then apply specific financial education to remove the barriers. This seems to work better than general financial education.

From a policy perspective, this paper suggests that policy makers should take into account the individuals' biases when design policies, especially in rural areas where most people are less educated. In particular, policy makers can provide cheap financial education to overcome individual constraints and thus improve individual welfares.

From a methodological standpoint, this paper is among the first to use a laboratory experiment as an intervention in the field experiment.[41] We find that the laboratory experiment influenced the field behavior in our setting. By using laboratory experiments, researchers can explicitly manipulate more variables which are endogenous or are otherwise difficult to manipulate. For example, Malmendier and Nagel (2010) find that individuals who have experienced low stock-market returns throughout their lives so far are less likely to participate in the stock market. However,

[41] As far as we know, the method is similar to Carter *et al.* (2008) and Gaurav *et al.*(2011).

it is difficult to manipulate experience in order to influence individual decisions. In this paper, we use a laboratory experiment to simulate experiences and influence field behaviors. We hope to explore in future research whether this can apply to other settings.

## Chapter Three: An Experiment on Reference Points and Expectations

### 1. Introduction

Kahneman and Tversky's (1979) prospect theory is well documented in the economic and psychology literature. In this theory, the evaluation of an outcome is influenced by how it compares to a reference point, the degree of diminishing sensitivity, loss aversion and nonlinear probability weighting. What determines a reference point is an important and open question for discussion. The status quo is one candidate for the reference point, which implies that individuals are reluctant to give up things they currently possess. Alternatively, expectations are taken to be reference points (Koszegi and Rabin 2006, 2007), meaning that individuals are reluctant to fall short of their beliefs. These theories about reference point determination have different implications due to loss aversion below reference point.

In the theories of expectation-based reference points, there are two lines of literature. One line of literature is the determination of certainty-equivalent reference points in models of Disappointment Aversion (DA) (Bell 1985; Loomes and Sugden 1986; Gul 1991). In DA models, the reference point is modeled as the expected utility certainty equivalent of a gamble. The outcome is evaluated by comparing it to a fixed number which equals the expected utility certainty equivalent. Another line is the determination of stochastic reference distributions in the more recent models of Koszegi and Rabin (KR) (2006, 2007). In the KR model the reference point is the full distribution of expected outcomes. The outcome is evaluated by comparing it with each expected outcome and then integrating over the distribution of expected outcome.

This paper tests to what extent expectations and the status quo determine the reference point based on different theoretical implications. I conducted a controlled lab experiment in which I explicitly manipulated expectations and exogenously varied expectations in different groups. I first randomly split the sample into the control group and the treatment group. Then I sent information to these groups in an email 24 hours before the experiment. For the control group, the email said that they will receive a fixed amount of payment for the experiment. For the treatment groups, the email said that they will receive a lottery as payment. When the subjects were in the lab, the treatment groups would play a lottery. Then both the control group and the treatment group would answer 60 risk-attitude questions to elicit their risk attitudes following the Holt and Laury (2002) procedure. The difference between the two groups will help to identify the role of expectations and the status quo as reference point. In particular, I explicitly manipulated expectations to be stochastic so that I could shed light on the distinctions between DA model and KR models.

I also explored the second source of variation: I exogenously varied the time of receiving new information and tested whether individuals assimilated new information into their reference points, and if so, at what rate. I randomly split the overall treatment group into two groups: the "no-waiting" treatment group and the "waiting" treatment group. The "no-waiting" group answered the questions immediately after they discovered the realization of the lottery. The "waiting" group filled out a survey about their social economic background after they knew the realization of the lottery, and then—after a few minutes—answered the questions. The key difference was that the waiting group risk attitudes were elicited five minutes later than the no-waiting group. This second source of variation further identified the role of expectation as reference point because the timing of new information does not influence the status quo. I also varied the payoffs and probabilities in the questions measuring risk attitudes so that I could use MLE to jointly estimate the reference points and the preferences based on the reference points.

I find that those who have higher expectations are less risk averse, and those who have lower expectations are more risk averse. The estimated reference points from MLE are higher in the group with higher expectations. These results suggest that expectations play a role to determine the reference point. I also find small diminishing sensitivity, significant loss aversion, and significant nonlinear probability weighting.

To investigate the relative importance of expectations and the status quo, I nested the two models and estimated the weight on each model. I find that the weight on expectation is 0.71, which suggest that both expectations and the status quo determine the reference point but expectations play a more important role.

To compare the model of the certainty-equivalent reference point with that of the stochastic reference point, I explicitly manipulated expectations to be stochastic. The structural estimation suggests that the model of the stochastic reference point fits my data better than that of the certainty-equivalent reference point.

I also exogenously varied the time of receiving new information and tested whether individuals adjust new information into their reference points and the speed of the adjustment. I nested the model of full adjustment and that of no adjustment, and estimated the weight on each model. The weight on the model of full adjustment is 0.54, which suggests that subjects adjust reference points quickly.

My work contributes to the literature in the following ways. First, it provides evidence on the question "What determines the reference point?" I test to what extent expectations and the status quo play a role. Many previous empirical research assumed the status quo, lagged status quo as the reference point or treated reference points as latent variables in different contexts.[42] In some recent research, reference points are treated as expectations in the context of taxi drivers labor supply (Doran 2007; Crawford and Meng 2011), large stake risky choices (Post et al. 2008), insurance choices (Barseghyan *et al.* 2011), professional golf (Pope and Schweitzer 2011) and competition in a real effort sequential-move tournament (Gill and Prowse 2010). This paper differs from the above in that I exogenously manipulate expectations and expectations are induced to be stochastic.

There are related laboratory experiments that also explicitly manipulate subjects' expectations, and then check whether this manipulation influences their effort

[42] For example, Tversky and Kahneman (1992), and Tanaka et al. (2010) assume the status quo as reference point in their lab experiments. Reference points are assumed to be lagged status quo (purchase price) for small investors (Odean 1998) and for homeowners (Genesove and Mayer 2001). In the literature of negative elasticity of labor supply and income targeting, most research treated reference points as latent variables (Camerer et al. 1997; Farber 2005; Fehr and Götte 2007; Farber 2008).

provision (Abeler *et al.* 2011) or valuation for some products (Smith 2008; Ericson and Fuster 2010; Heffetz and List 2012). The findings of most papers are consistent with reference-dependent models and support the notion that reference points are expectations. There is one exception: Heffetz and List (2012) manipulate subjects' expectations about owning a product and find the endowment effect. But the effect is unlikely to be due to expectations as reference points. This paper is similar with regard to the manipulation, but differs in the following aspects: first, this paper studies whether the manipulation of expectations influences risk attitudes, not effort provision or valuation. Second, expectations were manipulated as stochastic, and reference points were modeled as certainty equivalent and stochastic in structural estimation. Moreover, I exogenously varied the time of receiving new information, which further identifies the role of expectation as reference point.

Second, my research sheds light on the different versions of expectations-based reference-dependent models. Sprenger (2010) uses the inconsistency of utility elicitation between probability and certainty equivalent methodology to distinguish the DA model from the KR model. In my experiment, I explicitly manipulated expectations to be stochastic and elicited risk attitudes after the expectations have been realized, shattered by unfavorable outcomes, or surpassed by favorable outcomes. The detailed choice data helped me to estimate the certainty-equivalent reference points and the weights on the stochastic reference points as well as the preferences based on the reference points. Therefore, my research deepens our understanding about different expectations-based models.

As far as I know, this paper is among the first to jointly estimate the reference points and other parameters in utility functions. In the previous research estimating structural parameters in reference-dependent models, reference points are either assumed to be the status quo (Tversky and Kahneman 1992; Tanaka *et al.* 2010) or in a Preferred Personal Equilibrium (Sprenger 2010) or in a Choice-acclimating Personal Equilibrium (Barseghyan *et al.* 2011).[43] Rabin and Weizsäcker (2009) jointly estimate the reference points and other preference parameters in reference-dependent models. This paper is similar with regard to joint estimation, but differs in that I manipulated the stochastic expectations and varied the time of receiving new information. Since individuals might not be in personal equilibrium, the estimation should be closer to the reality.

Third, this research provides evidence on the speed of adjustment of the reference point by exogenously varying the time of receiving new information. Slow adjustment can generate lower risk aversion after losses and after gains. Post *et al.* (2008) specified a lagged function for adjustment of the reference point, and

[43] Koszegi and Rabin (2006, 2007) present these rational expectations equilibrium concepts. The Unacclimating Personal Equilibrium (UPE) is the personal equilibrium in which individuals' choices correspond to expectations. The Preferred Personal Equilibrium (PPE) is the UPE with the highest ex-ante expected utility. The Choice-acclimating Personal Equilibrium (CPE) is the personal equilibrium in which individuals' choices correspond to expectations and the choices are committed well in advance of the resolution of uncertainty.

estimated the influence of initial expectations and recent outcomes. Their results suggest that reference points tend to stick to earlier values; this effect is stronger for recent outcomes than for initial expectations. Gill and Prowse (2010) estimate the adjustment of reference points in a real effort sequential-move tournament. They find that reference points of second mover adjust to their own effort choice quickly, which is consistent with Choice-acclimating Personal Equilibrium. This paper differs in that I not only estimated the speed of adjustment of reference points, but also exogenously varied the time of receiving new information. This second source of variation further identified the role of expectation as reference point and adds more evidence on the adjustment of reference points.

The rest of the paper proceeds as follows. In Section 2, I describe the experimental design. In Section 3, I show the theory model and its implications under different assumptions of reference points. The main empirical results are discussed in Section 4. I estimate the structural model, including the certainty-equivalent reference points and the weights on the stochastic reference points as well as the preferences based on the reference points. I conclude in Section 5.

## 2. Experiment Design

The timeline of the experiment is described in Figure 1:

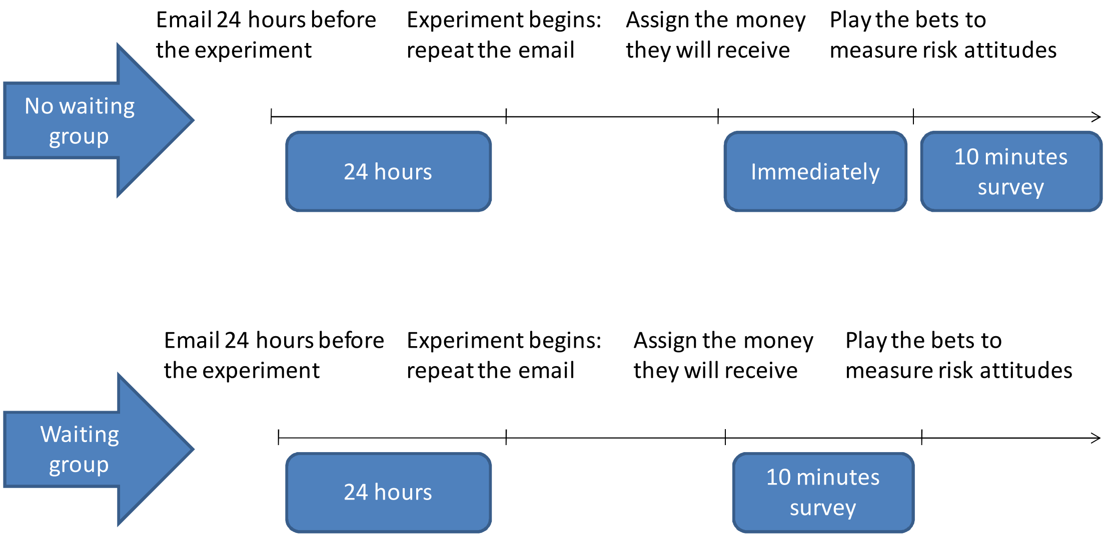



Figure 1 Timeline

Note: This figure shows the timeline for the no-waiting group and the waiting group

I first randomly split the sample into a control group, "no-waiting" treatment group, and "waiting" treatment group. Then I sent information to these groups in an email 24 hours before the experiment. Control group is randomly split into three subgroups: the \$10 control group, the \$15 control group and the \$20 control group. For the \$10 control group, the email says, "During the experiment, you will finish a short survey. After the survey, you will receive \$10." For the \$15 and \$20 control groups, the email says similar information except the amount they are going to receive. For the treatment groups, the email said, "During the experiment, you will finish a short survey. After the survey, you have 1/3 chance to receive \$10, 1/3 chance to receive \$15, and 1/3 chance to receive \$20." When the subjects were in the lab, the treatment groups would play a lottery and then ascertain whether they would receive $10, $15, or $20. Then both the control group and the two treatment groups would answer 60 risk-attitude questions to elicit their risk attitudes following Holt and Laury (2002) procedure (discussed below). In order to test whether subjects would adjust new information into their reference point and how quickly, the overall treatment group was split into two groups: "no-waiting" and "waiting."

As described in Figure 1, the "no-waiting" group answers the questions immediately after they discover whether they will receive $10, $15, or $20. The "waiting" group fills out a survey about their social economic background after they know whether they will get $10, $15, or $20, and then—after a few minutes—answers the questions. The key difference is that the waiting group risk attitudes are elicited five minutes later than the no-waiting group.

My design allows me to split all the subjects into 9 groups in the following table.

**Table 1. Summary of Lotter-Choice Treatments**

| | Control group | "No waiting" treatment | "Waiting" treatment |
|---|---|---|---|
| Comparison 1 | People who receive $10 from Control 1 | People who receive $10 from lottery (loser) | People who receive $10 from lottery (loser) |
| Comparison 2 | People who receive $15 from Control 2 | People who receive $15 from lottery | People who receive $15 from lottery |
| Comparison 3 | People who receive $20 from Control 3 | People who receive $20 from lottery (winner) | People who receive $20 from lottery (winner) |

This approach allowed me to undertake three comparisons about the subjects' risk attitudes. For example, in comparison 1, I compared the risk attitudes among people who receive $10 from Control 1; people who receive $10 from the lottery in the no-waiting treatment; and people who receive $10 from the lottery in the waiting treatment. Since the only differences among the groups were expectations at the time of choice and how long ago (24 hours vs 5 minutes) they were formed, I was able to test whether and how expectations matter.

For comparison 1, I use Tables A1, A2, and A3 to elicit risk attitudes. For comparisons 2 and 3, the tables are similar (see Appendix). The subjects are told that one of the bet outcomes will be randomly chosen ex post, so that they will report their true risk preference. In the experiment, subjects choose from option 1 and option 2 for each question. For table A1, when the probability of a high payoff increases (moving down the table), a subject should switch from option 1 (riskless option) to option 2 (risky option). The more riskless options the subject takes, the more risk averse the subject is. I use the number of riskless options taken as a measurement of risk aversion. The measurement from Table A1 is "measure 1". After Table A1, subjects answer one summary question in Table A4: "Now you have a choice between (1) Keep the $10 (2) Take the following bet: p% probability to get $15 and (100-p) % probability to get $5. What is the minimum probability p% that you will choose choice option 2?" "Measure 2" is calculated from this summary question. For example, if p=52, then measure 2 is 9 because the subject would take 9 riskless options if he/she answers the questions in Table 1.

The subject then answers the questions in Table A2 and Table A3. The measurement from Table A2 is "measure 3". After Table A2 and Table A3, subjects answer similar summary questions in Table A4. "Measure 4" and "measure 6" are calculated from these summary questions.

These tables differ in the following way. In Table A1, I fix the payoff but change the probability in the risky options. In Table A2, I fix the probability but change the payoff in the risky options. In Table A3, I fix the risky options but change the riskless options. There are summary questions after Tables A1, A2 and A3. The purpose is to have several measures of subjects' risk attitudes so that I can check the robustness of the results.

### 3. Theoretical Framework and Predictions

This section analyzes the predictions of the interventions if expectations determine reference points. In Cumulative Prospect Theory, I employ the specification from Post et al. (2008):

$$u(x \mid RP) = \begin{cases} (x - RP)^{\alpha} & if \quad x \geq RP \\ -\lambda (RP - x)^{\alpha} & if \quad x < RP \end{cases} \tag{1}$$

$\lambda > 0$ is the loss-aversion parameter[44], RP is the reference point that separates losses from gains, and $\alpha > 0$ measures the curvature of the value function, i.e. diminishing sensitivity.[45]

I also consider the one-parameter form of Drazen Prelec's (1998) axiomatically derived weighting function: $\pi(p) = \exp(-(-\ln p)^{\gamma})$. The probability weighting function is linear if $\gamma = 1$, as it is in EU. If $\gamma < 1$, the weighting function is inverted S-shaped, i.e., individuals overweight small probabilities and underweight large probabilities, as shown by Tversky and Kahneman (1992). If $\gamma > 1$, then the weighting function is S-shaped, i.e., individuals underweight small probabilities and overweight large probabilities.

There are three properties in the utility function: diminishing sensitivity, loss aversion and nonlinear probability weighting. These properties have the following implications on risk attitudes elicited from Table A1, A2 and A3. More diminishing sensitivity implies more risk aversion in the gain domain and more risk seeking in the loss domain. More loss aversion implies more risk aversion around the reference point. More nonlinear probability weighting implies more risk aversion in the gain domain

---

44 Koszegi and Rabin (2006, 2007) assume that overall utility has two components: consumption utility and gain-loss utility. They also assume $\eta > 0$ to be the weight the consumer attaches to gain-loss utility and $\lambda_{KR} > 1$ to be coefficient of loss aversion in gain-loss utility. Since $\eta$ is not identifiable, I just use gain-loss utility in my specification and the estimated $\lambda$ is $\frac{1 + \lambda_{KR}\eta}{1 + \eta}$ under KR's assumptions and the reference point r is zero.

45 The original formulation of prospect theory allows for different curvature parameters for the domain of losses and the domain of gains. To reduce the number of free parameters, I assume here that the curvature is equal for both domains.

and more risk seeking in the loss domain.

All three properties imply more risk seeking in the loss domain than around the kink. Diminishing sensitivity and nonlinear probability weighting imply more risk averse in the gain domain than around the kink but loss aversion implies less risk averse in the gain domain than around the kink. Therefore, I have hypothesis below.

***Hypothesis 1:*** Those whose reference points are greater than baseline will be less risk averse.

This hypothesis can be tested in comparison 1 when the reference points of subjects in the treatment group are greater than those in the $10 control group. The predication is that those in the treatment group are less risk averse.

***Hypothesis 2:*** Reference points adjust to the latest information about payoff.

This implies that the risk attitudes of the waiting group should be similar to those of the control group.

I do not have a clear prediction in comparison 3 when the reference points of subjects in the treatment group are less than those in the $20 control group. Those whose reference points are less than baseline might be less risk averse, more risk averse, or equally risk averse. If the effects of diminishing sensitivity and nonlinear probability weighting on risk attitudes are greater than that of los aversion, those in the treatment group are more risk averse. Otherwise, those in the treatment group are less risk averse.

## 4. Empirical Results

The experiment was conducted at the Experimental Social Science Laboratory (Xlab) at the University of California, Berkeley. The subjects in the experiment were recruited from undergraduate students. Each experimental session lasted about half an hour. Payoffs were calculated in dollars and the earnings were paid in private at the end of the experimental session.

A total of 396 subjects signed up for the experiments and received emails, and 306 of them actually showed up in 17 sections (see Table A5 for detail). Table 2 presents the summary statistics of the experiment.

**Table 2. Summary of Lotter-Choice Treatments**

| | Number of sessions | Number of subjects | Average earnings (USD) | Standard deviation |
|---|---|---|---|---|
| Control 1 | 3 | 47 | 10.16 | 3.6 |
| Control 3 | 3 | 47 | 19.98 | 2.03 |
| "No-waiting" treatment group | 6 | 125 | 14.32 | 5.16 |
| "Waiting" treatment group | 5 | 87 | 13.57 | 4.97 |
| Total | 17 | 306 | | |

The non-show up rate was 22.7%.[46] The non-show up rates for the $10 control

[46] According to the Xlab administrator, the average show up rate of Xlab experiments is about 65%-75%, and the show up rate of my experiment was a little higher than average.

group, the treatment group, and the $20 control group were 19.0%, 23.7%, and 21.7%, respectively. The Wald test shows that I cannot reject the equality of non-show up rates across different groups (p=0.70).

Post surveys show that 84% of subjects expect to get money in the range of my manipulations. In the $10 control group, 66% of subjects expected to get $10, 26% expected to get money between $10 and $15. In the $20 control group, 51% of subjects expected to get $20, 38% expected to get money between $10 and $20. In the treatment group, 95% of subjects expected to get money between $10 and $20. Therefore, subjects expected to get slightly more than my manipulation in the $10 control group and less than my manipulation in the $20 control group. Thus, the effects of expectation from my estimation are likely to be lower bounds.

The comparison between six measures can help to check individual consistency. According to the design, "measure 1" should be equal to "measure 2" because they are from equivalent questions. 46.4% of total subjects have the same "measure 1" and "measure 2". For 77.5% of subjects, the difference between "measure 1" and "measure 2" is no more than 2. The patterns are similar in other measures.[47] These results suggest that the measurements of risk attitudes are consistent cross different measurements.

### 4.2 The Effects of Expectations on Risk Attitudes

The main results are described in the following figures:

[47] "Measure 3" should be equal to "measure 4" and "measure 5" should be equal to "measure 6". 57.8% of total subjects have the same "measure 3" and "measure 4". For 76.5% of subjects, the difference between "measure 3" and "measure 4" is no more than 2. 63.6% of total subjects have the same "measure 5" and "measure 6". For 83.6% of subjects, the difference between "measure 5" and "measure 6" is no more than 2.

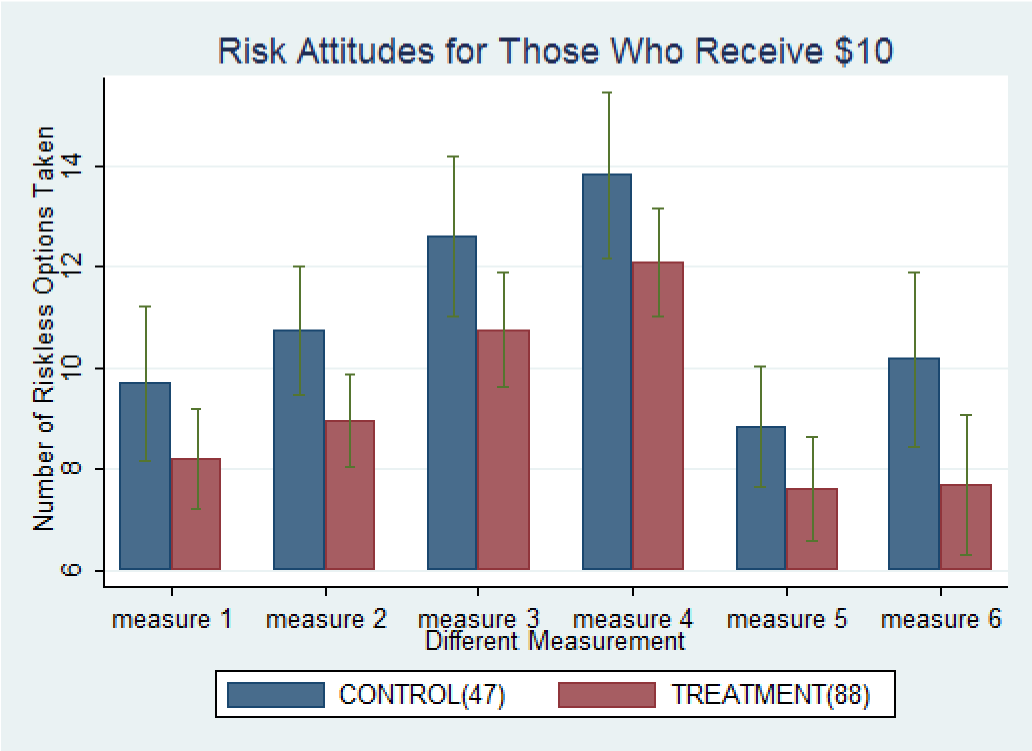



Figure 2 Risk attitudes for those who receive $10

Note: The figure shows the risk attitudes for those who receive $10 in the control group and the treatment groups. The vertical axis stands for the number of riskless options taken in Holt and Laury table and it measures risk averse. The six bars in the control group and the treatment group stand for different measures from the tables. "Measure 1" is derived from Table A1 that fixes payoffs but changes probability in risky options. "Measure 2" is calculated from the question (for comparison 1) "Now you have a choice between (1) Keep the $10 (2) Take the following bet: p% probability to get $15 and (100-p) % probability to get $5. What is the minimum probability p% that you will choose choice 2?" For example, if p=52, then measure 2 is 9 because the subject would take 9 riskless options if he/she answer the questions in Table 1. "Measure 3" is derived from Table A2 that fixes the probability to 50%/50% but change payoffs in risky options. "Measure 4" is similar to measure 2 but calculated from the question (for comparison 1) "Now you have a choice between (1) Keep the $10 (2) Take the following bet: 50% probability to get $X and 50% probability to get $5. What is the minimum X that you will choose choice 2?". "Measure 5" is derived from Table A3 that fixes the risky options but change the riskless options. "Measure 6" is similar to measure 2 and 4, but calculated from the question (for comparison 1) "Now you have a choice between (1) Keep $X (2) Take the following bet: 50% probability to get $10 and 50% probability to get $20. What is the minimum X that you will choose choice 1?"

Figure 2 shows the risk attitudes for those who received $10 in the control group and the treatment groups. The vertical axis stands for the number of riskless options taken in the Holt and Laury tables, and it measures risk aversion. The six bars in the control group and the treatment group stand for different measures from the tables. A risk neutral subject should take 9 riskless options in Table A1, 11 riskless options in Table A2 and 9 riskless options in Table A3. In the control group, the average of riskless options taken is 9.70 for Table A1, 12.60 for Table A2 and 8.83 for Table A3. Thus subjects are risk neutral or slightly risk averse in the control group. In the treatment group, the average of riskless options taken is 8.20 for Table A1, 10.75 for Table A2 and 7.62 for Table A3. This suggests that subjects are slightly risk seeking or risk neutral in the treatment group. The results from the above figure show a clear and consistent pattern that losers in the treatment groups are more risk seeking than those in the control group. This is consistent with the theory of reference-dependent utility with expectations as reference points.

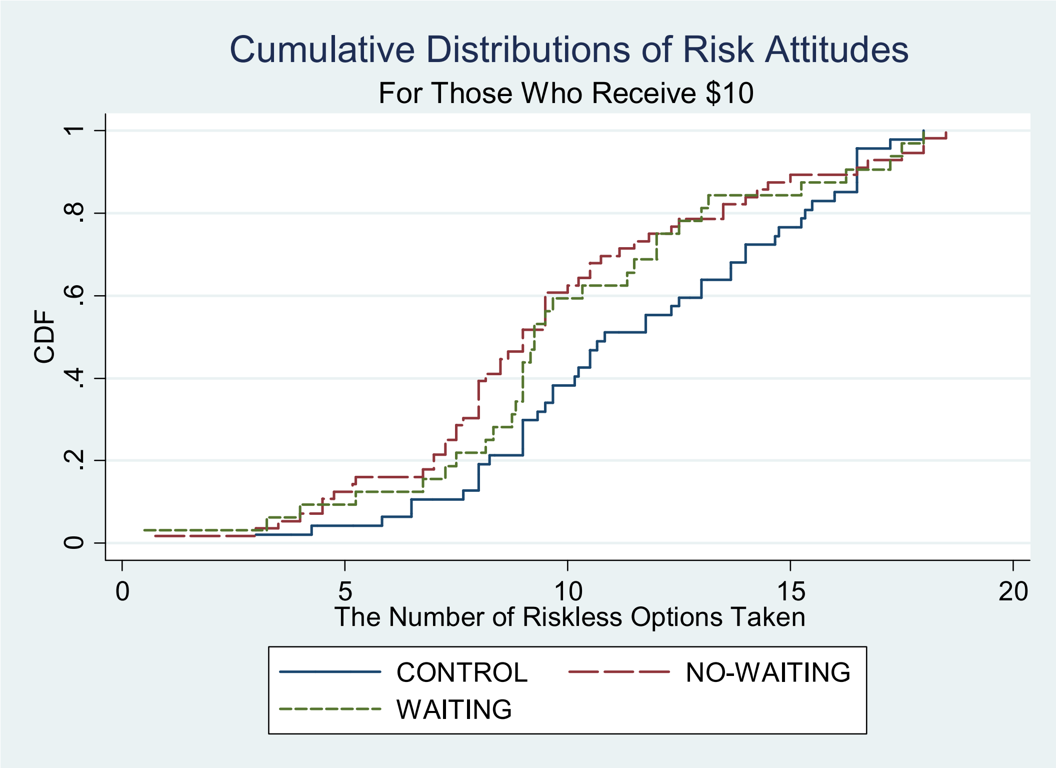



Figure 3 Cumulative distributions of risk attitudes for those who receive $10

Note: The figure shows the cumulative distributions of risk attitudes for those who receive $10 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury tables.

Figure 3 shows the cumulative distributions of risk attitudes for those who receive $10 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury tables. The figure shows that the risk aversion of the control group has first-order stochastic dominance over that of the no-waiting group. Kolmogorov-Smirnov test rejects the hypothesis that the distribution of risk attitudes in the control group is equal to that in the no-waiting group and it is significant at the 5% level. This is also consistent with the theory of reference-dependent utility with prospect theory value function and expectations as reference points.

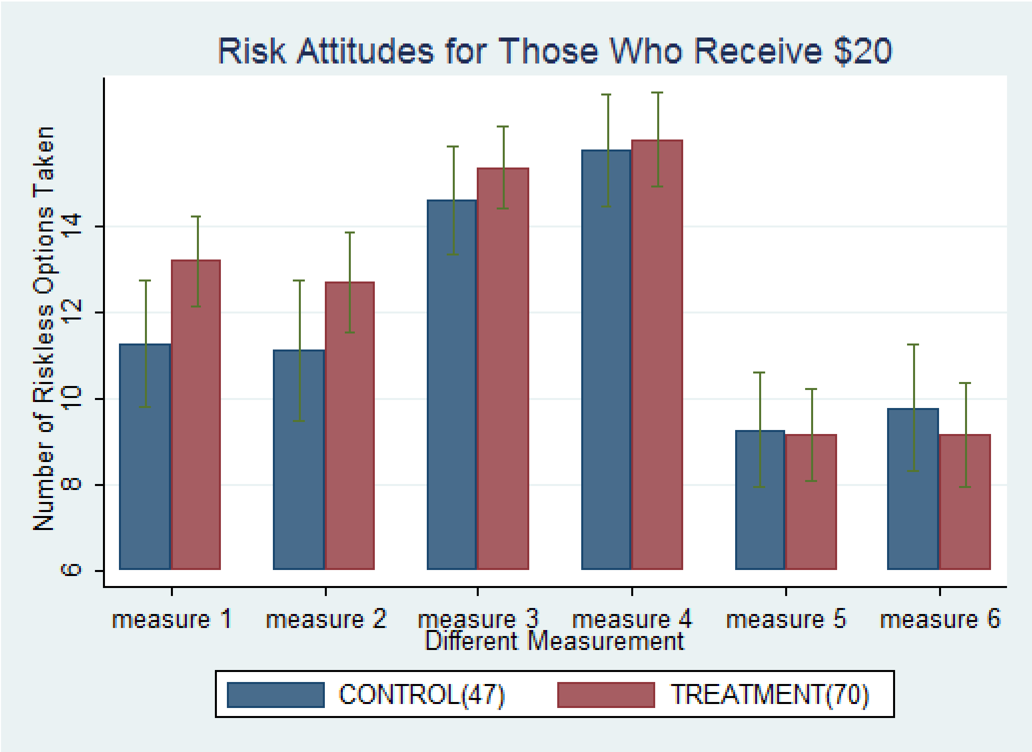



Figure 4 Risk attitudes for those who receive $20

Note: The figure shows the risk attitudes for those who receive $20 in the control group and the treatment groups. The vertical axis stands for the number of riskless options taken in Holt and Laury table and it measures risk averse. The six bars in the control group and the treatment group stand for different measures from the tables. The detail is explained in the note from figure 2

Figure 4 shows the risk attitudes for those who received $20 in the control group and the treatment groups. A risk neutral subject should take 9 riskless options in Table A1, 10 riskless options in Table A2 and 9 riskless options in Table A3. In the control group, the average of riskless options taken is 11.26 for Table A1, 14.60 for Table A2 and 9.26 for Table A3. Thus subjects are risk averse in the control group. In the treatment group, the average of riskless options taken is 13.17 for Table A1, 15.34 for Table A2 and 9.14 for Table A3. This suggests that subjects are also risk averse in the treatment group. The figure shows that winners in the treatment groups are more risk averse than those in the control group.

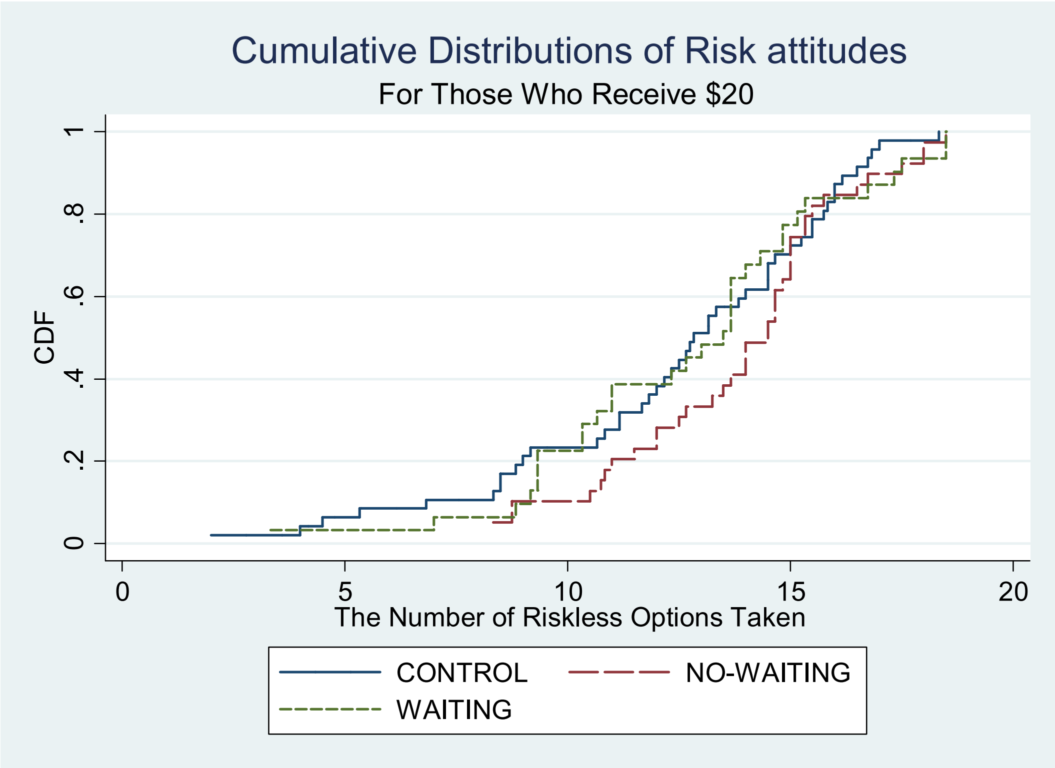



Figure 5 Cumulative distributions of risk attitudes for those who receive $20

Note: The figure shows the cumulative distributions of risk attitudes for those who receive $20 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury table.

Figure 5 shows the cumulative distributions of risk attitudes for those who receive $20 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury tables. The figure shows that the risk aversion of the no-waiting group has first-order stochastic dominance over that of the control group. Kolmogorov-Smirnov test cannot reject the hypothesis that the distribution of risk attitudes in the no-waiting group is equal to that in the control group (p=0.189).

In order to take into account other controls, I estimate the treatment effect of expectation on risk attitudes through OLS regression. For comparisons 1 and 3, I use the following specification:

$$y_i = \alpha + \beta_1 T_{nowait\_i} + \beta_2 T_{wait\_i} + \phi X_i + \varepsilon_i \qquad (2)$$

where $y_i$ is the number of riskless options taken by subject $i$. $T_{nowait\_i}$ is an indicator for no-waiting treatment and $T_{wait\_i}$ is an indicator for waiting treatment. $X_i$ is other control variables such as age and gender. $\beta_1$ captures the treatment effect of the expectations in the no-waiting group. $\beta_2$ captures the treatment effect of the expectations in the waiting group. $\beta_1 - \beta_2$ captures the effect of waiting on risk attitudes after the realization is revealed. $\varepsilon_i$ is assumed to be i.i.d. error term. I focus my analysis on two subsamples: those who receive $10 from the control group, the no-waiting group, or the waiting group (comparison 1 in Table 1), and those who receive $20 from the control group, the no-waiting group, or the waiting group (comparison 3 in Table 1). The results for six measures are presented in six columns in Tables 3 and 4.

**Table 3. The Effect of Expectation on Risk attitudes**

| **Specification:** | OLS | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | Number of riskless options taken | | | | | |
| **Sample:** | Those who receive $10 from control, no waiting group or waiting group | | | | | |
| | Measure 1 | Measure 2 | Measure 3 | Measure 4 | Measure 5 | Measure 6 |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| No waiting group with stochastic expectations | -2.18 | -1.80 | -2.13 | -1.79 | -0.94 | -2.98 |
| | (1.01)** | (0.87)** | (1.10)* | (1.04)* | (0.87) | (1.15)** |
| Waiting group with stochastic expectations | -0.87 | -2.30 | -1.64 | -2.21 | -1.68 | -2.39 |
| | (1.10) | (1.01)** | (1.22) | (1.19)* | (1.21) | (1.76) |
| Male | -0.02 | 0.22 | 0.33 | 0.04 | 0.46 | -1.24 |
| | (0.93) | (0.84) | (1.04) | (0.97) | (1.13) | (1.28) |
| Year in College | 0.41 | 0.24 | 0.32 | 0.27 | 0.27 | 0.33 |
| | (0.37) | (0.32) | (0.42) | (0.39) | (0.33) | (0.36) |
| Chi-square test: no | | | | | | |
| p-value | 0.2565 | 0.6375 | 0.7054 | 0.7144 | 0.5575 | 0.7397 |
| Omitted group | Those who receive $10 from the control group | | | | | |
| Obs. | 135 | 135 | 135 | 135 | 44 | 44 |
| Wave & time | Y | Y | Y | Y | N | N |
| R-square | 0.0492 | 0.0655 | 0.0399 | 0.0896 | 0.919 | 0.1425 |

Notes: Dependent variable is the number of riskless options taken measured by different tables; standard errors are clustered by each individual. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. Columns 1 to 6 report the results of measures 1 to 6, respectively.

Table 3 presents the results of comparison 1 in Table 1. The coefficient of $T_{nowait\_i}$ for measure 1 is -2.18 and is significant at the 5% level. So the losers in the no-waiting group choose about 2.18 fewer riskless options in Table A1 than those who receive $10 from the control group. It is the same pattern for other measures. The coefficients of $T_{nowait\_i}$ are negative for all six measures. And they are significant at the 10% level except for measure 5. So these results show a clear and consistent pattern that the losers from the no-waiting group are less risk averse (more risk seeking) than those who receive $10 from the control group. This is consistent with Hypothesis 1.

The coefficient of $T_{wait\_i}$ for measure 1 is -0.87 and is not significant. The magnitude of $\beta_2$ (0.87) is smaller than $\beta_1$ (2.18). The possible explanation is that the losers adjust their reference points to the realized payoffs ($10) and thus become more risk averse. However, the p-value of the Ward test $\beta_1 = \beta_2$ is 0.2565 and is not significant. There is thus suggestive evidence that losers in the waiting group are more risk averse than those in the no-waiting group with measure 1. In measures 2 to 6, $\beta_2$ is similar to $\beta_1$. So the risk attitudes of losers in the waiting group are similar to those in the no-waiting group. The reason might be that the subjects had to first finish the questions about measure 1, and then answer the questions about other measures. So there is a time lag between the realization of their lottery and the answers to the questions after measure 1. They had "waited" as if they were in the waiting group.

**Table 4. The Effect of Expectation on Risk attitudes**

| **Specification:** | OLS | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | Number of riskless options taken | | | | | |
| **Sample:** | Those who receive $20 from control, no waiting group or waiting group | | | | | |
| | Measure 1 | Measure 2 | Measure 3 | Measure 4 | Measure 5 | Measure 6 |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| No waiting group with stochastic expectations | 2.52 | 3.38 | 1.66 | 0.28 | -0.40 | -1.10 |
| | (0.93)*** | (1.04)*** | (0.86)* | (1.06) | (1.32) | (1.58) |
| Waiting group with stochastic expectations | 2.07 | 1.17 | 0.13 | -0.12 | 0.11 | -0.01 |
| | (1.16)* | (1.39) | (1.04) | (1.00) | (0.93) | (0.88) |
| Male | -0.76 | -0.90 | -1.33 | -0.73 | -1.04 | -0.76 |
| | (0.86) | (1.01) | (0.77)* | (0.89) | (0.76) | (0.77) |
| Year in College | 0.69 | 0.07 | 0.50 | 0.30 | 0.48 | 0.36 |
| | (0.38)* | (0.57) | (0.37) | (0.46) | (0.35) | (0.45) |
| Chi-square test: no waiting=waiting | | | | | | |
| p-value | 0.6976 | 0.1215 | 0.1553 | 0.7343 | 0.7137 | 0.5151 |
| Omitted group | Those who receive $20 from the control group | | | | | |
| Obs. | 117 | 117 | 117 | 117 | 78 | 78 |
| Wave & time | Y | Y | Y | Y | N | N |
| R-square | 0.1262 | 0.1216 | 0.0969 | 0.0796 | 0.0928 | 0.1174 |

Notes: Dependent variable is the number of riskless options taken measured by different tables; standard errors are clustered by each individual. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. Columns 1 to 6 report the results of measures 1 to 6, respectively.

Table 4 presents the results of comparison 3 in Table 1. The coefficient of $T_{nowait\_i}$ for measure 1 is 2.52 and is significant at the 1% level. It is the same pattern for measures 1 to 3. So the winners from the no-waiting group are more risk averse than those who receive $20 from the control group. This suggests that the effects of diminishing sensitivity and nonlinear probability weighting on risk attitudes are greater than that of loss aversion, Section 4.2 will estimate these preference parameters.

### 4.2 Structural Estimation

I have so far learned that expectations influence risk attitudes. This is consistent with expectations-based reference-dependent models. But how expectations might change the utility function is not discussed. There are at least two ways: the reference point could be modeled as a fixed number, which is the expected utility certainty-equivalent. Then outcome is evaluated by comparing it to a fixed number which equals the expected utility certainty equivalent. The reference point could also be modeled as the full distribution of expected outcomes (KR model). Then outcome is evaluated by comparing it with each expected outcome and then integrating over the distribution of expected outcome. In this section, I estimate these two models to deepen our understanding about different expectations-based models. In particular, I use the detailed choice data to estimate the certainty-equivalent reference points and the weights on the stochastic reference points as well as the preferences based on the reference points.

I can provide insights into the identification of these parameters. The lottery choice task identifies the utility function parameters. The subjects' choices made in Table A3 are used to estimate the curvature of utility function, since I fix the risky options but change the riskless options for all above exogenous reference points. The loss-aversion parameter is estimated using Table 2, since I fix the probabilities to 50%/50% but change the payoffs in risky options. The probability weighting parameter is estimated using Table A1, since I fix the payoffs but change the probabilities in risky options in that table. The experimental manipulation of expectations identifies the reference points in different groups.

To estimate the parameters in the utility function, I use a random-utility model (McFadden 1974) with a nonlinear component:

$$\tilde{u}(x) = \frac{1}{\sigma}u(x) + \varepsilon == \begin{cases} \frac{1}{\sigma}(x - RP)^{\alpha} + \varepsilon & if \quad x \geq RP \\ -\frac{1}{\sigma}\lambda(RP - x)^{\alpha} + \varepsilon & if \quad x < RP \end{cases} \tag{3}$$

where ε is assumed to be i.i.d. error term and modeled as type I extreme value. The utility is is scaled by 1/σ and the parameter σ is the scale parameter, because it scales the utility to reflect the variance of the unobserved portion of utility.

Suppose the subject is asked to choose between (1) Keep $x_1$ and (2) take the following bet: p probability to get $x_2$ and (1-p) probability to get $x_3$. Let $U(a)$ denote the utility as a function of their choices of bets. $a = 1$ if the subject chooses riskless options (option 1) and $a = 2$ if the subject chooses risky options (option 2). The probability to choose risky options can be presented by the usual logit formula:

$$P(a = 2) = \frac{\exp(U(a = 2))}{\exp(U(a = 1)) + \exp(U(a = 2))} \tag{4}$$

With this formula and the data about subjects' choices, I could use maximum-likelihood estimation to estimate the parameters in the structural model. Given that the underlying logistic model becomes highly nonlinear in the parameters, I code my own estimator in Stata to estimate the parameters and account for potential correlations within clusters. Below, I consider two expectations-based models: the model of the certainty-equivalent reference points and the model of the stochastic reference points.

#### 4.2.1 The Model of the Certainty-Equivalent Reference Points

According to equation (3), I will estimate $\alpha, \lambda, \gamma, \sigma$ and certainty-equivalent reference points. The parameters $\alpha, \lambda, \gamma, \sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha > 1$, $\lambda < 1$ and $\gamma > 1$. The log-likelihood is calculated through equation (4), where $U(a=1) = u(x_1)$ and $U(a=2) = pu(x_2) + (1-p)u(x_3)$ and p is the probability in the risky options. I also allow the coefficients of certainty-equivalent reference points in the following six groups to be different from each other: those who receive \$10 in the control group (rp1), those in the no-waiting treatment (rp2), those who receive \$10 in the waiting treatment (rp3), those who receive \$15 in the waiting treatment (rp4), those who receive \$20 in the waiting treatment (rp5), and those who receive \$20 in the control group (rp6). I present the results in Table 5.

**Table 5. Maximum Likelihood Estimation of Utility Function**

| **Model:** | **DA model with the fixed reference point** | | | | | |
|---|---|---|---|---|---|---|
| **Constraint:** | No constraint | Linear probability weighting | No diminishing sensitivity | No loss aversion | α-CPE | Expectation vs status quo |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| α | 1.18 | 0.85 | | 0.86 | 0.93 | 0.99 |
| | (0.05) | (0.09) | | (0.03) | (0.04) | (0.06) |
| λ | 1.86 | 2.05 | 0.0000 | | 1.56 | 1.25 |
| | (0.26) | (0.30) | (0.0000) | | (0.07) | (0.11) |
| σ | 1.72 | 1.39 | 1.01 | 0.71 | 0.97 | 1.04 |
| | (0.21) | (0.40) | (0.05) | (0.05) | (0.08) | (0.13) |
| γ | 0.49 | | 0.54 | 0.49 | 0.42 | 0.51 |
| | (0.04) | | (0.03) | (0.04) | (0.03) | (0.06) |
| Weight on the new outcome for no-waiting group (aa1) | | | | | 0.48 | |
| | | | | | (0.08) | |
| Weight on the new outcome for waiting group (aa2) | | | | | 0.63 | |
| | | | | | (0.10) | |
| Weight on expectations-based model (aa) | | | | | | 0.64 |
| | | | | | | (0.07) |
| Fixed reference point for $10 control group (rp1) | 2.78 | 5.40 | 5.08 | 9.13 | | |
| | (1.39) | (0.76) | (0.44) | (0.56) | | |
| Fixed reference point for no-waiting group (rp2) | 5.01 | 15.86 | 6.10 | 10.03 | | |
| | (0.05) | (1.03) | (0.35) | (0.21) | | |
| Fixed reference point for those who receive $10 in waiting group(rp3) | 7.32 | 36.77 | 5.64 | 9.75 | | |
| | (1.04) | (35.50) | (0.44) | (0.34) | | |
| Fixed reference point for those who receive $15 in waiting group(rp4) | 4.92 | 15.38 | 7.00 | 12.60 | | |
| | (2.10) | (0.67) | (0.07) | (1.06) | | |
| Fixed reference point for those who receive $20 in waiting group(rp5) | 20.23 | 19.48 | 10.08 | 15.82 | | |
| | (0.17) | (0.65) | (2.24) | (0.89) | | |
| Fixed reference point for $20 control group (rp6) | 19.87 | 19.08 | 11.13 | 16.55 | | |
| | (0.61) | (0.98) | (1.59) | (0.28) | | |
| Z test: λ=1 | | | | | | |
| p-value | 0.000*** | 0.000*** | | | 0.000*** | 0.021** |
| Z test: weight aa=0.5 | | | | | | |
| p-value | | | | | | 0.044** |
| Chi-square test: rp1=rp2 | | | | | | |
| p-value | 0.0022*** | | | 0.1370 | | |
| No of individuals | 306 | 306 | 306 | 306 | 306 | 306 |
| No of observation | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 |
| Log likelihood | -6326 | -6647 | -6446 | -6426 | -6463 | -6446 |

Note: I will estimate $\alpha, \lambda, \gamma, \sigma$ and certainty-equivalent reference points in utility function

$$\tilde{u}(x) = \begin{cases} \frac{1}{\sigma}(x-RP)^{\alpha} & if \quad x \geq RP \\ -\frac{1}{\sigma}\lambda(RP-x)^{\alpha} & if \quad x < RP \end{cases}.$$

through MLE. I apply 32 initials values in MLE and present the maximum of log-likelihood in the 32 estimations. The parameters $\alpha, \lambda, \gamma, \sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha > 1$, $\lambda < 1$ and $\gamma > 1$. I also allow the coefficients of certainty-equivalent reference points in the following six groups to be different from each other: those who receive \$10 in the control group (rp1), those in the no-waiting treatment (rp2), those who receive \$10 in the waiting treatment (rp3), those who receive \$15 in the waiting treatment (rp4), those who receive \$20 in the waiting treatment (rp5), and those who receive \$20 in the control group (rp6). Column 1 presents the estimation results with all seven parameters. In column 2, I constrain $\gamma = 1$ and estimate the model with linear probability weighting. In column 3, I constrain $\alpha = 1$ and estimate the model with no diminishing sensitivity. In column 4, I constrain $\lambda = 1$ and estimate the model with no loss aversion. In column 5, I estimate the weight on the utility from new reference points. I also allow the weights to be different in the no-waiting group (aa1) and the waiting group (aa2). In column 6, I compare the model of certainty-equivalent expectations as reference points with that of the status quo as reference points. I estimate the weight on the first model.

Column 1 presents the estimation results with all seven parameters. The point estimate of $\alpha$ is 1.18 and it is significantly greater than one at the 1% level. This is not consistent with diminishing sensitivity. The point estimate of $\lambda$ is 1.86 and it is significantly greater than one at the 1% level. This value of loss aversion is consistent with loss aversion estimates from other contexts (Tversky and Kahneman 1992; Gill and Prowse 2010; Pope and Schweitzer 2011). The point estimate of $\gamma$ is 0.49 and it is significantly less than one at the 1% level. The value of $\gamma$ is lower than estimates from other contexts, which is close to 0.7 (Tanaka *et al.* 2010; Barseghyan *et al.* 2010). The estimated certainty-equivalent reference points have the following pattern: $rp1 < rp2 < rp4$. The difference between rp1 and rp2 is significant at the 1% level. The difference between rp2 and rp4 is significant at the 1% level. The results are consistent with my interventions that give different expectations to different groups.

In column 2, I constrain $\gamma = 1$ and estimate the model with linear probability weighting. In column 3, I constrain $\alpha = 1$ and estimate the model with no diminishing sensitivity. In column 4, I constrain $\lambda = 1$ and estimate the model with no loss aversion. Then I can use standard likelihood ratio tests to investigate whether the model in column 1 fits significantly better than the models in columns 2 to 4. The test statistic D has a chi-squared distribution with k degrees of freedom:

$$\mathrm{D} = -2\,[\log \mathrm{L}(constrained) - \log \mathrm{L}(unconstrained)]$$

where $\mathrm{L}(constrained)$ is the likelihood of the constrained model, $\mathrm{L}(unconstrained)$ is the likelihood of the unconstrained model, and k is the difference in the number of degrees of freedom between the two models. The likelihood ratio tests show that the model in column 1 fits significantly better than the models in columns 2, 3, and 4 at the 1% level. These results suggest that probability weighting, curvature of utility function, and loss aversion all play an important role when the reference point is modeled as expected utility certainty equivalent.

In column 5, I analyze how fast the certainty-equivalent reference points adjust. I assume subjects will put some weight on the utility from new reference points and the rest on the utility from old reference points. The subjects do not adjust reference points if the weight is zero and they fully adjust reference points to new ones if the weight is one. I allow the weights to be different in the no-waiting group (aa1) and the waiting group (aa2). The point estimate of aa1 is 0.48 and of aa2 is 0.63, and both are significantly greater than zero at the 1% level. The positive weight in the no-waiting group suggests that subjects adjust reference points fast. The difference between them is not significant. Thus it is suggestive evidence that the longer they wait, the more they adjust their reference points to updated expectations.

In column 6, I compare the model of certainty-equivalent expectations as reference points with that of the status quo as reference points. I construct two utility functions: In the first function, the reference points are the certainty-equivalent expectations, i.e., rp1=10, rp2 =rp3=rp4=rp5=15, and rp6=20; in the second one, the reference points are the status quo, i.e. all reference points are zero. Then I estimate the weight on expectations. If the weight is larger than 0.5, the model with certainty-equivalent expectations as reference points is better than that with the status quo as reference points. The estimated weight is 0.64 and it is significantly greater than 0.5 at the 5% level. This result suggests that expectations are better than the status quo as reference points in the model of certainty-equivalent reference points.

### 4.2.2 The Model of the Stochastic Reference Points

In the KR model, the reference point is the full distribution of expected outcomes. For example, subjects expected to receive \$10 with 1/3 probability, \$15 with 1/3 probability and \$20 with 1/3 probability in the no-waiting group. Therefore, the reference points should be stochastic reference points with weights equal to the probability: \$10 with 1/3 probability, \$15 with 1/3 probability and \$20 with 1/3 probability. In KR model, I will estimate $\alpha, \lambda, \gamma, \sigma$ and the weights on stochastic reference points. The parameters $\alpha, \lambda, \gamma, \sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha > 1$, $\lambda < 1$ and $\gamma > 1$. The log-likelihood is calculated through equation (4), where

$$U(a=1) = w_{i1} \cdot u(x_1 \mid 10) + w_{i2} \cdot u(x_1 \mid 15) + w_{i3} \cdot u(x_1 \mid 20)$$

$$\begin{aligned} U(a=2) &= \pi(p) \cdot \{w_{i1} \cdot u(x_2 \mid 10) + w_{i2} \cdot u(x_2 \mid 15) + w_{i3} \cdot u(x_2 \mid 20)\} \\ &+ \pi(1-p) \cdot \{w_{i1} \cdot u(x_3 \mid 10) + w_{i2} \cdot u(x_3 \mid 15) + w_{i3} \cdot u(x_3 \mid 20)\} \end{aligned}$$

I allow the coefficients of the weights on stochastic reference points in the following six groups (i=1, …, 6) to be different from each other: those who receive \$10 in the control group (w11, w12 and w13), those in the no-waiting treatment (w21, w22 and w23), those who receive \$10 in the waiting treatment (w31, w32 and w33), those who receive \$15 in the waiting treatment (w41, w42 and w43), those who receive \$20 in the waiting treatment (w51, w52 and w53), and those who receive \$20 in the control group (w61, w62 and w63). Since $w_{i1} + w_{i2} + w_{i3} = 1$, there are 11 parameters to estimate ($\alpha, \lambda, \gamma, \sigma$ and 12 weights in four groups) in KR model and it is not identified.

Therefore, I constraint $w_{i2} = w_{i3}$ and estimate only 9 parameters. I present the results in column 1 Table 6.

**Table 6. Maximum Likelihood Estimation of Utility Function**

| **Model:** | **KR model with stochastic reference points** | | | | | | KR vs DA |
|---|---|---|---|---|---|---|---|
| **Constraint:** | No constraint | Linear probability weighting | No diminishing | No loss aversion | α-CPE | Expectation vs status | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| α | 1.00 | 0.72 | | 0.96 | 1.04 | 1.21 | 1.13 |
| | (0.07) | (0.10) | | (0.04) | (0.04) | (0.08) | (0.06) |
| λ | 2.06 | 2.77 | 2.05 | | 1.60 | 1.37 | 1.70 |
| | (0.27) | (0.28) | (0.26) | | (0.08) | (0.15) | (0.12) |
| σ | 1.18 | 1.31 | 1.17 | 0.86 | 1.19 | 1.78 | 1.66 |
| | (0.22) | (0.36) | (0.12) | (0.07) | (0.11) | (0.36) | (0.22) |
| γ | 0.37 | | 0.38 | 0.46 | 0.39 | 0.43 | 0.44 |
| | (0.03) | | (0.03) | (0.04) | (0.03) | (0.04) | (0.04) |
| Weight on the new outcome for no-waiting group (aa1) | | | | | 0.40 | | |
| | | | | | (0.09) | | |
| Weight on the new outcome for waiting group (aa2) | | | | | 0.54 | | |
| | | | | | (0.12) | | |
| Weight on expectations-based model (aa) | | | | | | 0.71 | |
| | | | | | | (0.06) | |
| Weight on KR model (aa) | | | | | | | 1.00 |
| | | | | | | | (0.00) |
| Weight for reference point as \$10 for \$10 control group(w11) | 0.65 | 0.41 | 0.65 | 1.00 | | | |
| | (0.07) | (0.08) | (0.07) | (0.0000) | | | |
| Weight for reference point as \$10 for no-waiting group(w21) | 0.51 | 0.16 | 0.51 | 1.00 | | | |
| | (0.08) | (0.11) | (0.07) | (0.0000) | | | |
| Weight for reference point as \$10 for those who receive \$10 in waiting group(w31) | 0.56 | 0.30 | 0.57 | 1.00 | | | |
| | (0.07) | (0.10) | (0.07) | (0.0000) | | | |
| Weight for reference point as \$10 for those who receive \$15 in waiting group(w41) | 0.90 | 0.35 | 0.89 | 1.00 | | | |
| | (0.23) | (0.24) | (0.24) | (0.0000) | | | |
| Weight for reference point as \$10 for those who receive \$20 in waiting group(w51) | 0.0000 | 0.0000 | 0.0000 | 0.65 | | | |
| | (0.0000) | (0.0000) | (0.0000) | (0.24) | | | |
| Weight for reference point as \$10 for \$20 control group(w61) | 0.0000 | 0.0000 | 0.0000 | 0.51 | | | |
| | (0.0000) | (0.0000) | (0.0000) | (0.21) | | | |
| Z test: λ=1 | | | | | | | |
| p-value | 0.000*** | 0.000*** | 0.000*** | | 0.000*** | 0.013** | |
| Z test: weight aa=0.5 | | | | | | | |
| p-value | | | | | | 0.000*** | |
| Chi-square test: w11=w21 | | | | | | | |
| p-value | 0.064* | | | | | | |
| No of individuals | 306 | 306 | 306 | 306 | 306 | 306 | 306 |
| No of observation | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 |
| Log likelihood | -6389 | -6610 | -6389 | -6501 | -6467 | -6415 | -6619 |

Note: I will estimate $\alpha, \lambda, \gamma, \sigma$ and certainty-equivalent reference points in utility function

$$\tilde{u}(x)=\begin{cases}\frac{1}{\sigma}(x-RP)^{\alpha} & if \quad x\geq RP\\ -\frac{1}{\sigma}\lambda(RP-x)^{\alpha} & if \quad x<RP\end{cases}$$

through MLE. I apply 32 initials values in MLE and present the maximum of log-likelihood in the 32 estimations. The parameters $\alpha,\lambda,\gamma,\sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha>1$, $\lambda<1$ and $\gamma>1$. Column 1 presents the estimation results with all seven parameters in the KR model. In column 2, I constrain $\gamma=1$ and estimate the model with linear probability weighting. In column 3, I constrain $\alpha=1$ and estimate the model with no diminishing sensitivity. In column 4, I constrain $\lambda=1$ and estimate the model with no loss aversion. In column 5, I estimate the weight on the utility from new reference points. I also allow the weights to be different in the no-waiting group (aa1) and the waiting group (aa2). In column 6, I compare the model of stochastic expectations as reference points with that of the status quo as reference points. I estimate the weight on the first model. In column 7, I compare the model of the stochastic reference point with that of the certainty-equivalent reference point

Column 1 presents the estimation results with all seven parameters in the KR model. The point estimate of $\alpha$ is 1.00 and it is not significantly different from one. The point estimate of $\lambda$ is 2.06 and it is significantly greater than one at the 1% level. This value of loss aversion is also consistent with loss aversion estimates from other contexts. The estimated weights on stochastic reference points have the following pattern: $w_{11}>w_{21}>w_{61}$. The difference between $w_{11}$ and $w_{21}$ is significant at the 10% level. The results are consistent with my interventions that give different expectations to different groups.

In columns 2, 3 and 4, I estimate the models with linear probability weighting, no diminishing sensitivity and no loss aversion, similar to Table 5. The likelihood ratio tests show that the model in column 1 fits significantly better than the models in columns 2 and 4 at the 1% level. However, column 1 and 3 are similar. These results suggest both nonlinear probability weighting and loss aversion play an important role in the KR model. But the curvature is close to linear.

In column 5, I analyze how fast the stochastic weights on reference points adjust in the similar way to Table 5. The point estimate of aa1 is 0.40 and of aa2 is 0.54 in the KR model, and both are significantly greater than zero at the 1% level. The positive weight in the no-waiting group suggests that subjects adjust reference points fast. The difference between them is not significant. Thus it is suggestive evidence that the longer they wait, the more they adjust their reference points to new ones.

In column 6, I compare the model of stochastic expectations as reference points with that of the status quo as reference points in the similar way to Table 5. The estimated weight on expectations is 0.71 and it is significantly greater than 0.5 at the 1% level. This result suggests that expectations are better than the status quo as reference points in the KR model.

In column 7, I compare the model of the stochastic reference point with that of the certainty-equivalent reference point. I construct two utility functions: In the first function, I use stochastic expectations as reference points, i.e., $w_{11}=1$, $w_{21}=w_{31}=w_{41}=w_{51}=\frac{1}{3}$, $w_{61}=0$; I use certainty-equivalent expectations as reference points, i.e. ,rp1=10, rp2 =rp3=rp4=rp5=15, and rp6=20. Then I estimate the weight on the first model. The estimated weight is 1.00 and it is significantly greater than 0.5 at the 1% level. Therefore, the model of the stochastic reference point fits my data better than that of the certainty-equivalent reference point.

### 4.3 Calibration

I have estimated the parameters and weights on reference points in Table 6. In this section, I will use the estimated parameters and weights to calibrate subjects' behavior in my experiment. Note I will only consider the control group and the no-waiting group in the calibration.

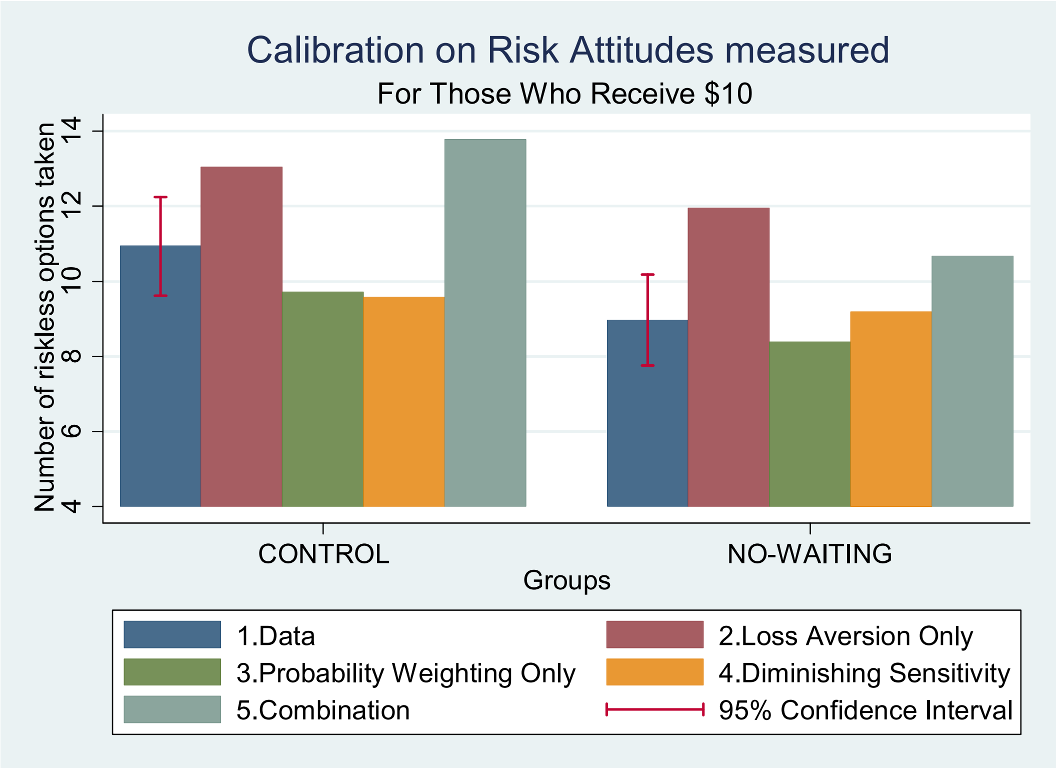



Figure 6

Note: I calibrate the risk attitudes for those who receive $10. I use the specification

$$u(x \mid RP) = \begin{cases} (x - RP)^{\alpha} & if \quad x \geq RP \\ -\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}$$

and the one-parameter form of probability weighting function: $\pi(p) = \exp(-(-\ln p)^{\gamma})$. The vertical axis represents the average number of riskless options taken from different measurements. In the control group, the reference points are $20 in the calibration. In the no-waiting group, the weights on stochastic reference points are 0.70 on $10, 0.15 on $15 and 0.15 on $20 in the calibration. There are five models in both the control group and the no-waiting group. The following is the list of models:

1st: Actual risk attitudes in my data

2nd: No diminishing sensitivity, loss aversion, linear probability weighting ($\alpha = \gamma = 1, \lambda = 1.81$)

3rd: No diminishing sensitivity, no loss aversion, nonlinear probability weighting ($\alpha = \lambda = 1, \gamma = 0.38$)

4th: Diminishing sensitivity, no loss aversion, linear probability weighting ($\lambda = \gamma = 1, \alpha = 0.92$)

5th: Diminishing sensitivity, loss aversion, nonlinear probability weighting ($\alpha = 0.92$, $\lambda = 1.81$ and $\gamma = 0.38$)

Figure 6 shows the calibration on risk attitudes for those who receive $10. Since the patterns are similar in different measures of risk attitudes, I report the average number of riskless options taken in the vertical axis.

This figure is consistent with the analysis in Section 3. Diminishing sensitivity, loss aversion and nonlinear probability weighting imply more risk seeking in the loss domain than around the kink. It also suggests that the results of my MLE estimation could generate the similar pattern of my reduced form results: in the 5th model combining all three properties, the difference between the control and the treatment has the right direction and the levels of risk attitudes fall in 95% confidence interval.

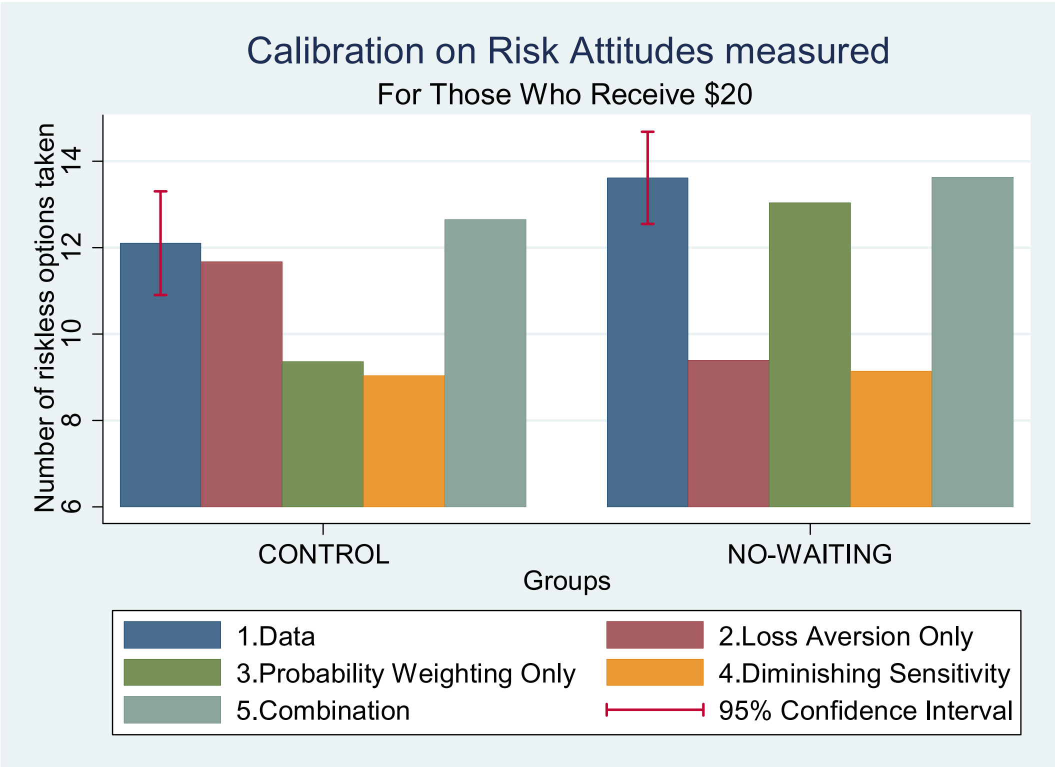



Figure 7

Note: I calibrate the risk attitudes for those who receive \$20. I use the specification

$$u(x \mid RP) = \begin{cases} (x - RP)^{\alpha} & if \quad x \geq RP \\ -\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}$$

and the one-parameter form of probability weighting function: $\pi(p) = \exp(-(-\ln p)^{\gamma})$. The vertical axis represents the average number of riskless options taken from different measurements. In the control group, the reference points are \$20 in the calibration. In the no-waiting group, the weights on stochastic reference points are 0.70 on \$10, 0.15 on \$15 and 0.15 on \$20 in the calibration. There are five models in both the control group and the no-waiting group. The following is the list of models:

1st: Actual risk attitudes in my data

2nd: No diminishing sensitivity, loss aversion, linear probability weighting ($\alpha = \gamma = 1, \lambda = 1.81$)

3rd: No diminishing sensitivity, no loss aversion, nonlinear probability weighting ($\alpha = \lambda = 1, \gamma = 0.38$)

4th: Diminishing sensitivity, no loss aversion, linear probability weighting ($\lambda = \gamma = 1, \alpha = 0.92$)

5th: Diminishing sensitivity, loss aversion, nonlinear probability weighting ($\alpha = 0.92$, $\lambda = 1.81$ and $\gamma = 0.38$)

Figure 7 shows the calibration on risk attitudes for those who receive \$20. Diminishing sensitivity and nonlinear probability weighting imply more risk averse in the gain domain than around the kink but loss aversion implies less risk averse in the gain domain than around the kink. It also suggests that the results of my MLE estimation could generate the similar pattern of my reduced form results: in the 5th model combining all three properties, the difference between the control and the treatment has the right direction and the levels of risk attitudes fall in 95% confidence interval.

## 5. Conclusion

What determines a reference point is an important question. This paper provides evidence whether expectations and the status quo determine the reference point. I explicitly manipulated expectations and exogenously varied expectations in different groups. Then I tested whether expectations change subjects' risk attitudes. I find that both expectations and the status quo determine the reference point but expectations play a more important role. Moreover, the structural estimation suggests that the model of the stochastic reference point fits my data better than that of the certainty-equivalent reference point.

I also exogenously varied the time of receiving new information and tested whether individuals adjust new information into their reference points, and the speed of the adjustment. I find that subjects can incorporate much new information into reference points in a few minutes, suggesting that subjects adjust reference points quickly. Prior work and this paper suggests that expectation, the status quo, the time of holding previous beliefs and the time of adjusting new information contribute to determine reference points together. Future work should tell these apart, with both field and laboratory evidence.

**References**


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Human (2011), "Reference Points and Effort Provision," *American Economic Review, 101(2):470-492.*

Barseghyan, Levon, Molinari, Francesca, O'Donoghue, Ted and Teitelbaum, Joshua C. (2011), "The Nature of Risk Preferences: Evidence from Insurance Choices". Available at SSRN: http://ssrn.com/abstract=1646520

Bell, David E., "Disappointment in Decision Making under Uncertainty," *Operations Research*, 1985, 33 (1), 1-27.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler (1997), "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics, 112 (2), 407-441.*

Crawford, Vincent P. and Juanjuan Meng (2011), "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review, 101(5): 1912–32.*

Ericson, Keith M. Marzilli and Andreas Fuster (2009), "Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior," Working Paper.

Farber(2005), H. "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers," *Journal of Political Econom*y, 113.

Farber, H. (2008) "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review, 98(2): 1069-1082*.

Fehr, E. and Goette, L. (2007) "Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment." *American Economic Review*, *97(1).*

Genesove, D. and C. Mayer (2001). "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics, 116(4), 1233-1260.*

Gul, Faruk (1991), "A Theory of Disappointment Aversion," *Econometrica*, 59 (3), 667-686.

Heffetz, Ori and John List (2012) "Assignment, Expectations, and Endowment: Variations in the Lab and Their Effects on Choice," *Working Paper*

Holt, Charles A. and Susan K. Laury (2002), "Risk Aversion and Incentive Effects," *The American Economic Review, 92 (5), 1644-1655.*

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, *47 (2), 263-291.*

Koszegi, Botond and Matthew Rabin (2006), "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, *121 (4), 1133-1165*.

- And (2007) "Reference-Dependent Risk Attitudes," *The American Economic Review*, *97 (4), 1047-1073.*
- And (2009), "Reference-Dependent Consumption Plans," *American Economic Review*, *2009, 99 (3), 909-936.*

Loomes, Graham and Robert Sugden (1986), "Disappointment and Dynamic Consistency in Choice under Uncertainty," *Review of Economic Studies*, *53 (2), 271-82.*

Odean, Terrance(1998), "Are Investors Reluctant to Realize Their Losses?" *The Journal of Finance*, *53 (5), 177-1798.*

Pope, Devin and Maurice Schweitzer (2011), "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *American Economic Review*, *101(1):129-57*

Post, Thierry, Martihn J. van den Assem, Guido Baltussen, and Richard H.Thaler (2008), “Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show," *American Economic Review*, *98 (1), 38-71.*

Prelec, Drazen (1998), “The Probability Weighting Function,” *Econometrica*, *66 (3), 497-527*

Samuelson, William and Richard Zeckhauser, “Status Quo Bias in Decision Making,” Journal of Risk and Uncertainty, 1 (1988), 7-59.

Sprenger, Charles, “An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Point”, UCSD

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen (2010), “Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam,” *American Economic Review, 100(1): 557–71*

Tversky, Amos and Craig R. Fox(1995), “Weighing Risk and Uncertainty," *Psychological Review*, *102 (2), 269-283.*

—and Daniel Kahneman(1992), “Advances in Prospect Theory: Cumulative Representation of Uncertainty," Journal of Risk and Uncertainty*, 5 (4), 297-323.*

**Table 1. Summary of Lotter-Choice Treatments**

| | Control group | “No waiting” treatment | “Waiting” treatment |
|---|---|---|---|
| Comparison 1 | People who receive $10 from Control 1 | People who receive $10 from lottery (loser) | People who receive $10 from lottery (loser) |
| Comparison 2 | People who receive $15 from Control 2 | People who receive $15 from lottery | People who receive $15 from lottery |
| Comparison 3 | People who receive $20 from Control 3 | People who receive $20 from lottery (winner) | People who receive $20 from lottery (winner) |

**Table 2. Summary of Lotter-Choice Treatments**

| | Number of sessions | Number of subjects | Average earnings (USD) | Standard deviation |
|---|---|---|---|---|
| Control 1 | 3 | 47 | 10.16 | 3.6 |
| Control 3 | 3 | 47 | 19.98 | 2.03 |
| "No-waiting" treatment group | 6 | 125 | 14.32 | 5.16 |
| "Waiting" treatment group | 5 | 87 | 13.57 | 4.97 |
| Total | 17 | 306 | | |

Table 3. The Effect of Expectation on Risk attitudes

| **Specification:** | OLS | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | Number of riskless options taken | | | | | |
| **Sample:** | Those who receive $10 from control, no waiting group or waiting group | | | | | |
| | Measure 1 | Measure 2 | Measure 3 | Measure 4 | Measure 5 | Measure 6 |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| No waiting group with stochastic expectations | -2.18 | -1.80 | -2.13 | -1.79 | -0.94 | -2.98 |
| | (1.01)** | (0.87)** | (1.10)* | (1.04)* | (0.87) | (1.15)** |
| Waiting group with stochastic expectations | -0.87 | -2.30 | -1.64 | -2.21 | -1.68 | -2.39 |
| | (1.10) | (1.01)** | (1.22) | (1.19)* | (1.21) | (1.76) |
| Male | -0.02 | 0.22 | 0.33 | 0.04 | 0.46 | -1.24 |
| | (0.93) | (0.84) | (1.04) | (0.97) | (1.13) | (1.28) |
| Year in College | 0.41 | 0.24 | 0.32 | 0.27 | 0.27 | 0.33 |
| | (0.37) | (0.32) | (0.42) | (0.39) | (0.33) | (0.36) |
| Chi-square test: no | | | | | | |
| p-value | 0.2565 | 0.6375 | 0.7054 | 0.7144 | 0.5575 | 0.7397 |
| Omitted group | Those who receive $10 from the control group | | | | | |
| Obs. | 135 | 135 | 135 | 135 | 44 | 44 |
| Wave & time | Y | Y | Y | Y | N | N |
| R-square | 0.0492 | 0.0655 | 0.0399 | 0.0896 | 0.919 | 0.1425 |

Notes: Dependent variable is the number of riskless options taken measured by different tables; standard errors are clustered by each individual. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. Columns 1 to 6 report the results of measures 1 to 6, respectively.

**Table 4. The Effect of Expectation on Risk attitudes**

| **Specification:** | OLS | | | | | |
|---|---|---|---|---|---|---|
| **Dep. Var.:** | Number of riskless options taken | | | | | |
| **Sample:** | Those who receive $20 from control, no waiting group or waiting group | | | | | |
| | Measure 1 | Measure 2 | Measure 3 | Measure 4 | Measure 5 | Measure 6 |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| No waiting group with stochastic expectations | 2.52 | 3.38 | 1.66 | 0.28 | -0.40 | -1.10 |
| | (0.93)*** | (1.04)*** | (0.86)* | (1.06) | (1.32) | (1.58) |
| Waiting group with stochastic expectations | 2.07 | 1.17 | 0.13 | -0.12 | 0.11 | -0.01 |
| | (1.16)* | (1.39) | (1.04) | (1.00) | (0.93) | (0.88) |
| Male | -0.76 | -0.90 | -1.33 | -0.73 | -1.04 | -0.76 |
| | (0.86) | (1.01) | (0.77)* | (0.89) | (0.76) | (0.77) |
| Year in College | 0.69 | 0.07 | 0.50 | 0.30 | 0.48 | 0.36 |
| | (0.38)* | (0.57) | (0.37) | (0.46) | (0.35) | (0.45) |
| Chi-square test: no waiting=waiting | | | | | | |
| p-value | 0.6976 | 0.1215 | 0.1553 | 0.7343 | 0.7137 | 0.5151 |
| Omitted group | Those who receive $20 from the control group | | | | | |
| Obs. | 117 | 117 | 117 | 117 | 78 | 78 |
| Wave & time | Y | Y | Y | Y | N | N |
| R-square | 0.1262 | 0.1216 | 0.0969 | 0.0796 | 0.0928 | 0.1174 |

Notes: Dependent variable is the number of riskless options taken measured by different tables; standard errors are clustered by each individual. Robust clustered standard errors are in the parentheses. *** significant on 1% level; ** significant on 5% level, * significant on 10% level. Columns 1 to 6 report the results of measures 1 to 6, respectively.

**Table 5. Maximum Likelihood Estimation of Utility Function**

| **Model:** | **DA model with the fixed reference point** | | | | | |
|---|---|---|---|---|---|---|
| **Constraint:** | No constraint | Linear probability weighting | No diminishing sensitivity | No loss aversion | α-CPE | Expectation vs status quo |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| α | 1.18 | 0.85 | | 0.86 | 0.93 | 0.99 |
| | (0.05) | (0.09) | | (0.03) | (0.04) | (0.06) |
| λ | 1.86 | 2.05 | 0.0000 | | 1.56 | 1.25 |
| | (0.26) | (0.30) | (0.0000) | | (0.07) | (0.11) |
| σ | 1.72 | 1.39 | 1.01 | 0.71 | 0.97 | 1.04 |
| | (0.21) | (0.40) | (0.05) | (0.05) | (0.08) | (0.13) |
| γ | 0.49 | | 0.54 | 0.49 | 0.42 | 0.51 |
| | (0.04) | | (0.03) | (0.04) | (0.03) | (0.06) |
| Weight on the new outcome for no-waiting group (aa1) | | | | | 0.48 | |
| | | | | | (0.08) | |
| Weight on the new outcome for waiting group (aa2) | | | | | 0.63 | |
| | | | | | (0.10) | |
| Weight on expectations-based model (aa) | | | | | | 0.64 |
| | | | | | | (0.07) |
| Fixed reference point for $10 control group (rp1) | 2.78 | 5.40 | 5.08 | 9.13 | | |
| | (1.39) | (0.76) | (0.44) | (0.56) | | |
| Fixed reference point for no-waiting group (rp2) | 5.01 | 15.86 | 6.10 | 10.03 | | |
| | (0.05) | (1.03) | (0.35) | (0.21) | | |
| Fixed reference point for those who receive $10 in waiting group(rp3) | 7.32 | 36.77 | 5.64 | 9.75 | | |
| | (1.04) | (35.50) | (0.44) | (0.34) | | |
| Fixed reference point for those who receive $15 in waiting group(rp4) | 4.92 | 15.38 | 7.00 | 12.60 | | |
| | (2.10) | (0.67) | (0.07) | (1.06) | | |
| Fixed reference point for those who receive $20 in waiting group(rp5) | 20.23 | 19.48 | 10.08 | 15.82 | | |
| | (0.17) | (0.65) | (2.24) | (0.89) | | |
| Fixed reference point for $20 control group (rp6) | 19.87 | 19.08 | 11.13 | 16.55 | | |
| | (0.61) | (0.98) | (1.59) | (0.28) | | |
| Z test: λ=1 | | | | | | |
| p-value | 0.000*** | 0.000*** | | | 0.000*** | 0.021** |
| Z test: weight aa=0.5 | | | | | | |
| p-value | | | | | | 0.044** |
| Chi-square test: rp1=rp2 | | | | | | |
| p-value | 0.0022*** | | | 0.1370 | | |
| No of individuals | 306 | 306 | 306 | 306 | 306 | 306 |
| No of observation | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 |
| Log likelihood | -6326 | -6647 | -6446 | -6426 | -6463 | -6446 |

Note: I will estimate $\alpha, \lambda, \gamma, \sigma$ and certainty-equivalent reference points in utility function

$$\tilde{u}(x) = \begin{cases} \frac{1}{\sigma}(x - RP)^{\alpha} & if \quad x \geq RP \\ -\frac{1}{\sigma}\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}.$$

through MLE. I apply 32 initials values in MLE and present the maximum of log-likelihood in the 32 estimations. The parameters $\alpha, \lambda, \gamma, \sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha > 1$, $\lambda < 1$ and $\gamma > 1$. I also allow the coefficients of certainty-equivalent reference points in the following six groups to be different from each other: those who receive \$10 in the control group (rp1), those in the no-waiting treatment (rp2), those who receive \$10 in the waiting treatment (rp3), those who receive \$15 in the waiting treatment (rp4), those who receive \$20 in the waiting treatment (rp5), and those who receive \$20 in the control group (rp6). Column 1 presents the estimation results with all seven parameters. In column 2, I constrain $\gamma = 1$ and estimate the model with linear probability weighting. In column 3, I constrain $\alpha = 1$ and estimate the model with no diminishing sensitivity. In column 4, I constrain $\lambda = 1$ and estimate the model with no loss aversion. In column 5, I estimate the weight on the utility from new reference points. I also allow the weights to be different in the no-waiting group (aa1) and the waiting group (aa2). In column 6, I compare the model of certainty-equivalent expectations as reference points with that of the status quo as reference points. I estimate the weight on the first model.

**Table 6. Maximum Likelihood Estimation of Utility Function**

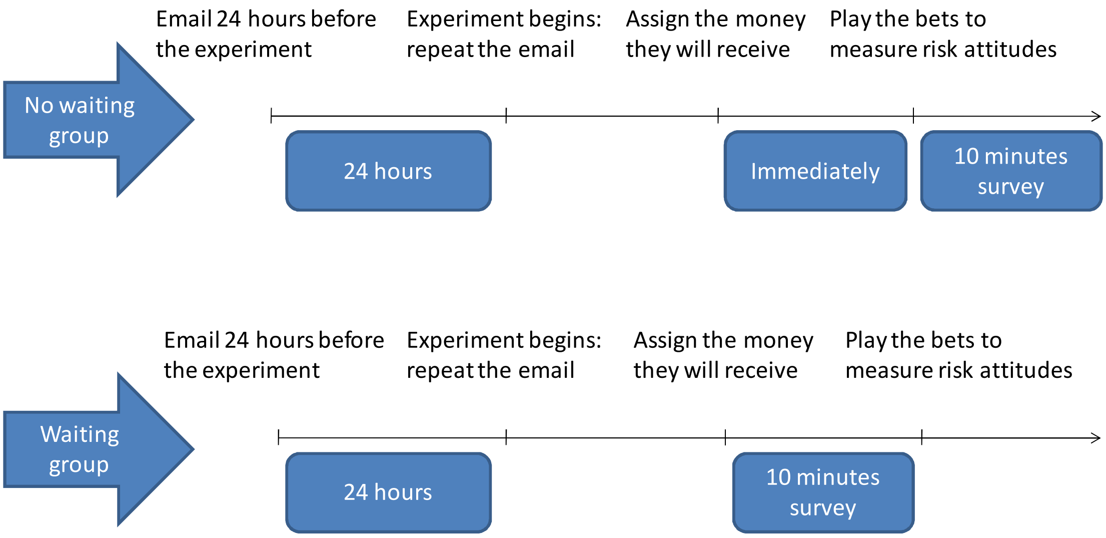
| **Model:** | **KR model with stochastic reference points** | | | | | | KR vs DA |
|---|---|---|---|---|---|---|---|
| **Constraint:** | No constraint | Linear probability weighting | No diminishing | No loss aversion | α-CPE | Expectation vs status | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| α | 1.00 | 0.72 | | 0.96 | 1.04 | 1.21 | 1.13 |
| | (0.07) | (0.10) | | (0.04) | (0.04) | (0.08) | (0.06) |
| λ | 2.06 | 2.77 | 2.05 | | 1.60 | 1.37 | 1.70 |
| | (0.27) | (0.28) | (0.26) | | (0.08) | (0.15) | (0.12) |
| σ | 1.18 | 1.31 | 1.17 | 0.86 | 1.19 | 1.78 | 1.66 |
| | (0.22) | (0.36) | (0.12) | (0.07) | (0.11) | (0.36) | (0.22) |
| γ | 0.37 | | 0.38 | 0.46 | 0.39 | 0.43 | 0.44 |
| | (0.03) | | (0.03) | (0.04) | (0.03) | (0.04) | (0.04) |
| Weight on the new outcome for no-waiting group (aa1) | | | | | 0.40 | | |
| | | | | | (0.09) | | |
| Weight on the new outcome for waiting group (aa2) | | | | | 0.54 | | |
| | | | | | (0.12) | | |
| Weight on expectations-based model (aa) | | | | | | 0.71 | |
| | | | | | | (0.06) | |
| Weight on KR model (aa) | | | | | | | 1.00 |
| | | | | | | | (0.00) |
| Weight for reference point as $10 for $10 control group(w11) | 0.65 | 0.41 | 0.65 | 1.00 | | | |
| | (0.07) | (0.08) | (0.07) | (0.0000) | | | |
| Weight for reference point as $10 for no-waiting group(w21) | 0.51 | 0.16 | 0.51 | 1.00 | | | |
| | (0.08) | (0.11) | (0.07) | (0.0000) | | | |
| Weight for reference point as $10 for those who receive $10 in waiting group(w31) | 0.56 | 0.30 | 0.57 | 1.00 | | | |
| | (0.07) | (0.10) | (0.07) | (0.0000) | | | |
| Weight for reference point as $10 for those who receive $15 in waiting group(w41) | 0.90 | 0.35 | 0.89 | 1.00 | | | |
| | (0.23) | (0.24) | (0.24) | (0.0000) | | | |
| Weight for reference point as $10 for those who receive $20 in waiting group(w51) | 0.0000 | 0.0000 | 0.0000 | 0.65 | | | |
| | (0.0000) | (0.0000) | (0.0000) | (0.24) | | | |
| Weight for reference point as $10 for $20 control group(w61) | 0.0000 | 0.0000 | 0.0000 | 0.51 | | | |
| | (0.0000) | (0.0000) | (0.0000) | (0.21) | | | |
| Z test: λ=1 | | | | | | | |
| p-value | 0.000*** | 0.000*** | 0.000*** | | 0.000*** | 0.013** | |
| Z test: weight aa=0.5 | | | | | | | |
| p-value | | | | | | 0.000*** | |
| Chi-square test: w11=w21 | | | | | | | |
| p-value | 0.064* | | | | | | |
| No of individuals | 306 | 306 | 306 | 306 | 306 | 306 | 306 |
| No of observation | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 | 14288 |
| Log likelihood | -6389 | -6610 | -6389 | -6501 | -6467 | -6415 | -6619 |

Note: I will estimate $\alpha, \lambda, \gamma, \sigma$ and certainty-equivalent reference points in utility function

$$\tilde{u}(x) = \begin{cases} \frac{1}{\sigma}(x - RP)^{\alpha} & if \quad x \geq RP \\ -\frac{1}{\sigma}\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}$$

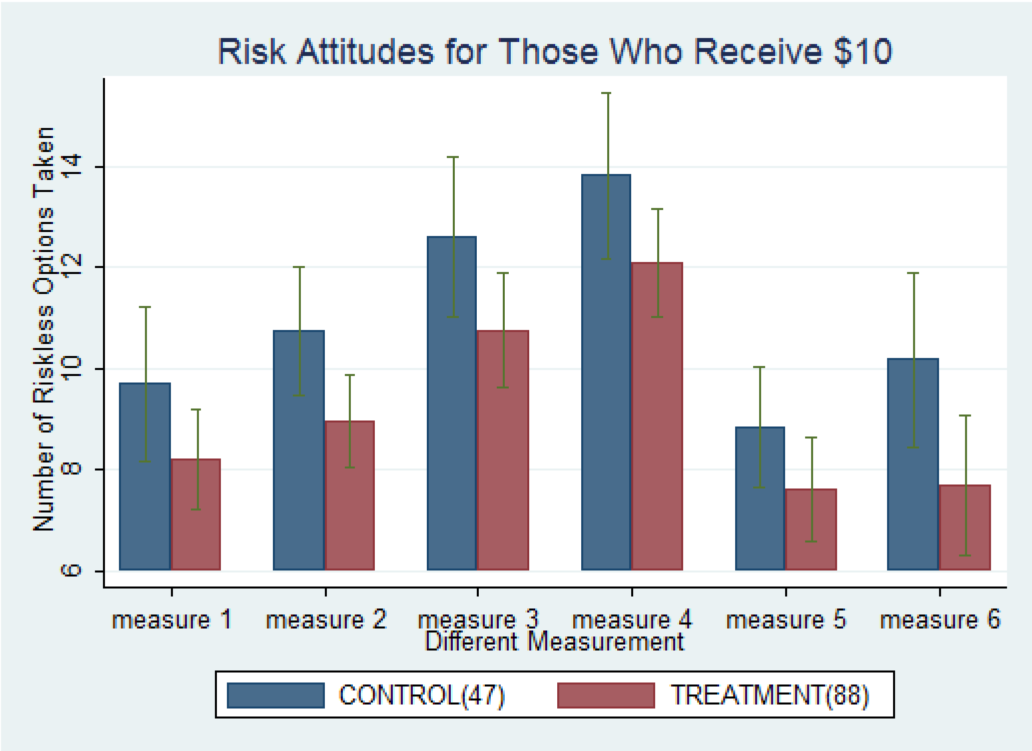
through MLE. I apply 32 initials values in MLE and present the maximum of log-likelihood in the 32 estimations. The parameters $\alpha, \lambda, \gamma, \sigma$ are constrained to be positive and thus I allow the possibilities of $\alpha > 1$, $\lambda < 1$ and $\gamma > 1$. Column 1 presents the estimation results with all seven parameters in the KR model. In column 2, I constrain $\gamma = 1$ and estimate the model with linear probability weighting. In column 3, I constrain $\alpha = 1$ and estimate the model with no diminishing sensitivity. In column 4, I constrain $\lambda = 1$ and estimate the model with no loss aversion. In column 5, I estimate the weight on the utility from new reference points. I also allow the weights to be different in the no-waiting group (aa1) and the waiting group (aa2). In column 6, I compare the model of stochastic expectations as reference points with that of the status quo as reference points. I estimate the weight on the first model. In column 7, I compare the model of the stochastic reference point with that of the certainty-equivalent reference point

Figure 1 Timeline



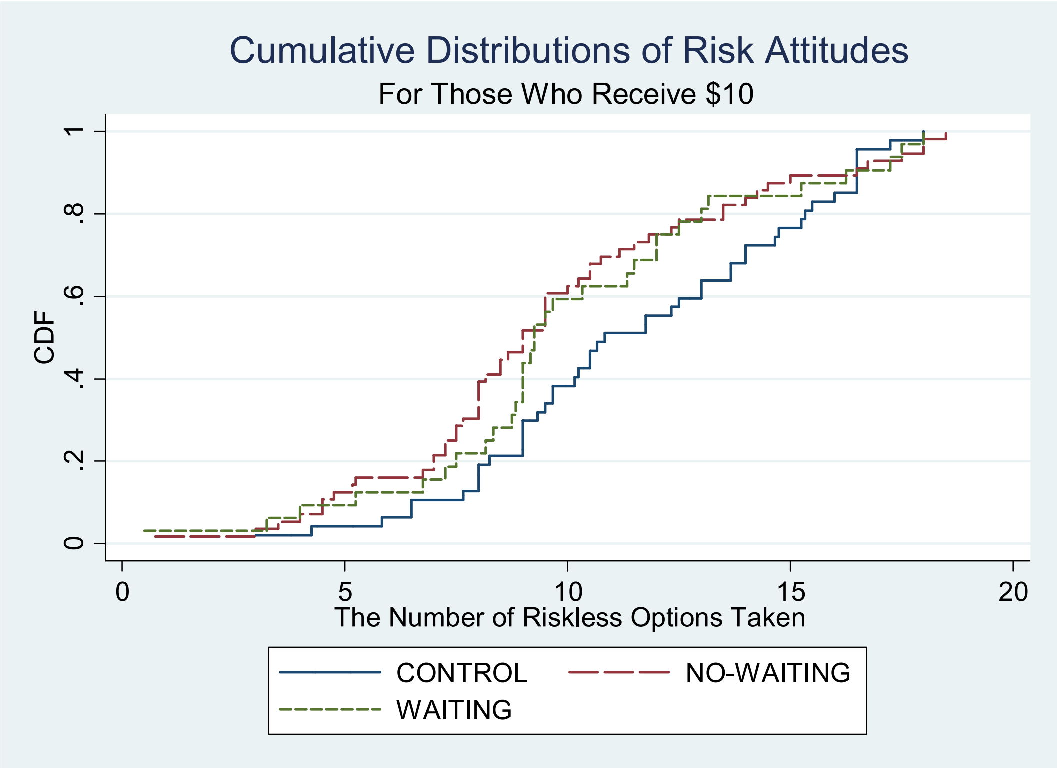

Note: This figure shows the timeline for the no-waiting group and the waiting group

Figure 2 Risk attitudes for those who receive $10



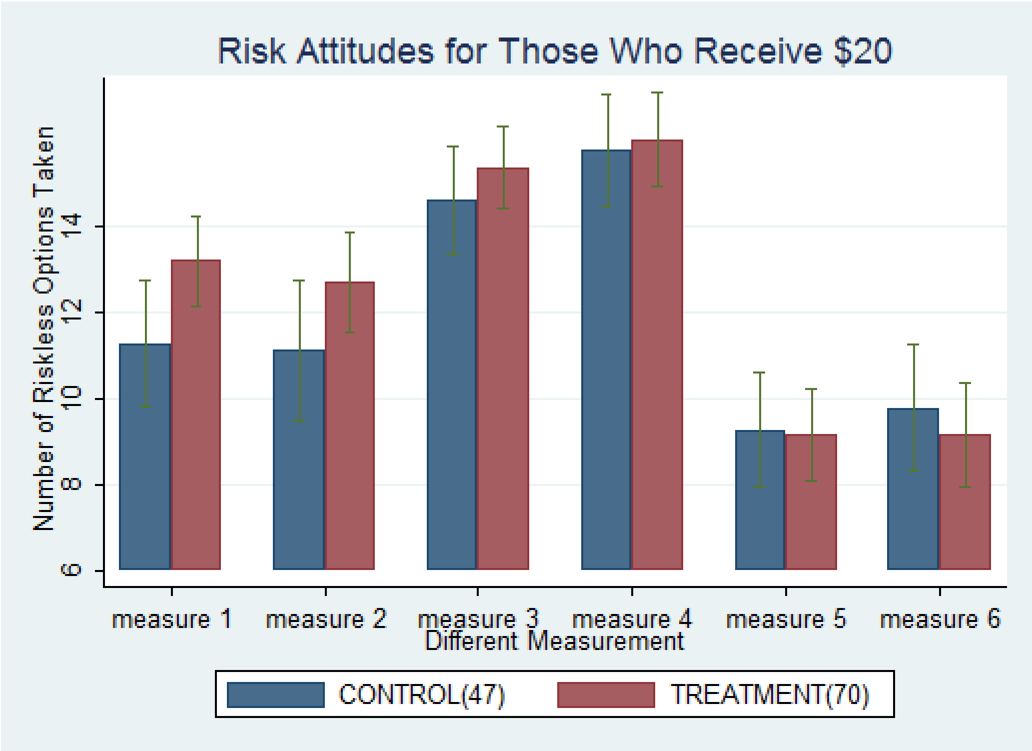

Note: The figure shows the risk attitudes for those who receive $10 in the control group and the treatment groups. The vertical axis stands for the number of riskless options taken in Holt and Laury table and it measures risk averse. The six bars in the control group and the treatment group stand for different measures from the tables. "Measure 1" is derived from Table A1 that fixes payoffs but changes probability in risky options. "Measure 2" is calculated from the question (for comparison 1) "Now you have a choice between (1) Keep the $10 (2) Take the following bet: p% probability to get $15 and (100-p) % probability to get $5. What is the minimum probability p% that you will choose choice 2?" For example, if p=52, then measure 2 is 9 because the subject would take 9 riskless options if he/she answer the questions in Table 1. "Measure 3" is derived from Table A2 that fixes the probability to 50%/50% but change payoffs in risky options. "Measure 4" is similar to measure 2 but calculated from the question (for comparison 1) "Now you have a choice between (1) Keep the $10 (2) Take the following bet: 50% probability to get $X and 50% probability to get $5. What is the minimum X that you will choose choice 2?". "Measure 5" is derived from Table A3 that fixes the risky options but change the riskless options. "Measure 6" is similar to measure 2 and 4, but calculated from the question (for comparison 1) "Now you have a choice between (1) Keep $X (2) Take the following bet: 50% probability to get $10 and 50% probability to get $20. What is the minimum X that you will choose choice 1?"

Figure 3 Cumulative distributions of risk attitudes for those who receive $10



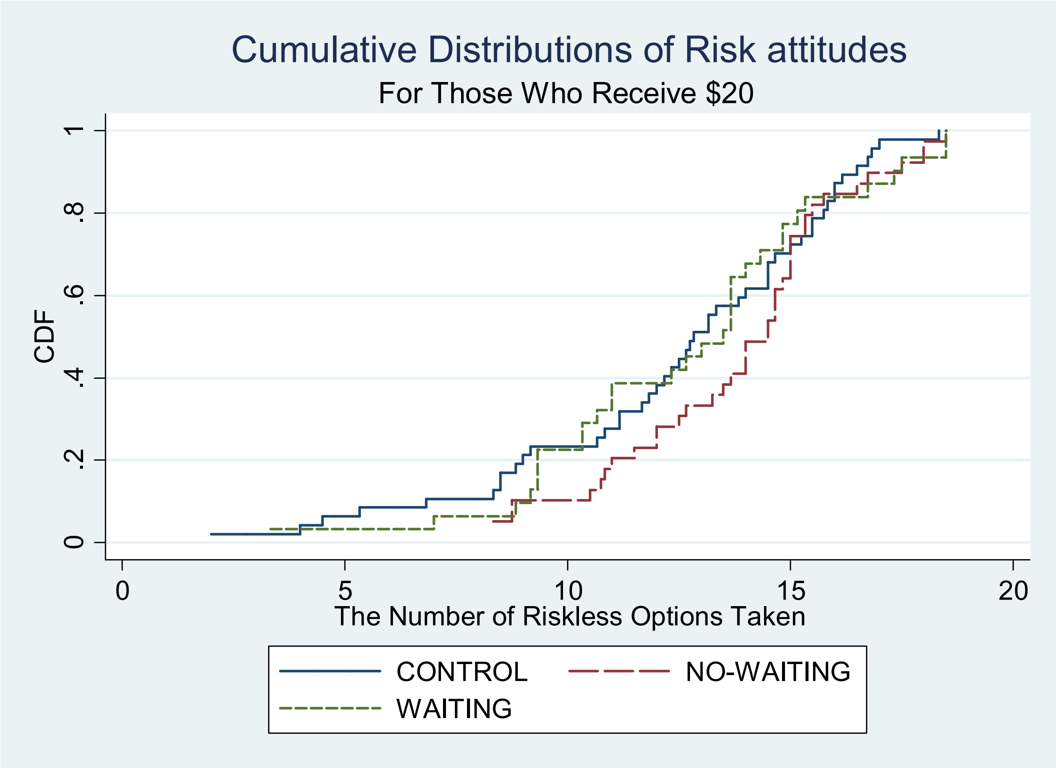

Note: The figure shows the cumulative distributions of risk attitudes for those who receive $10 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury tables.

Figure 4 Risk attitudes for those who receive $20

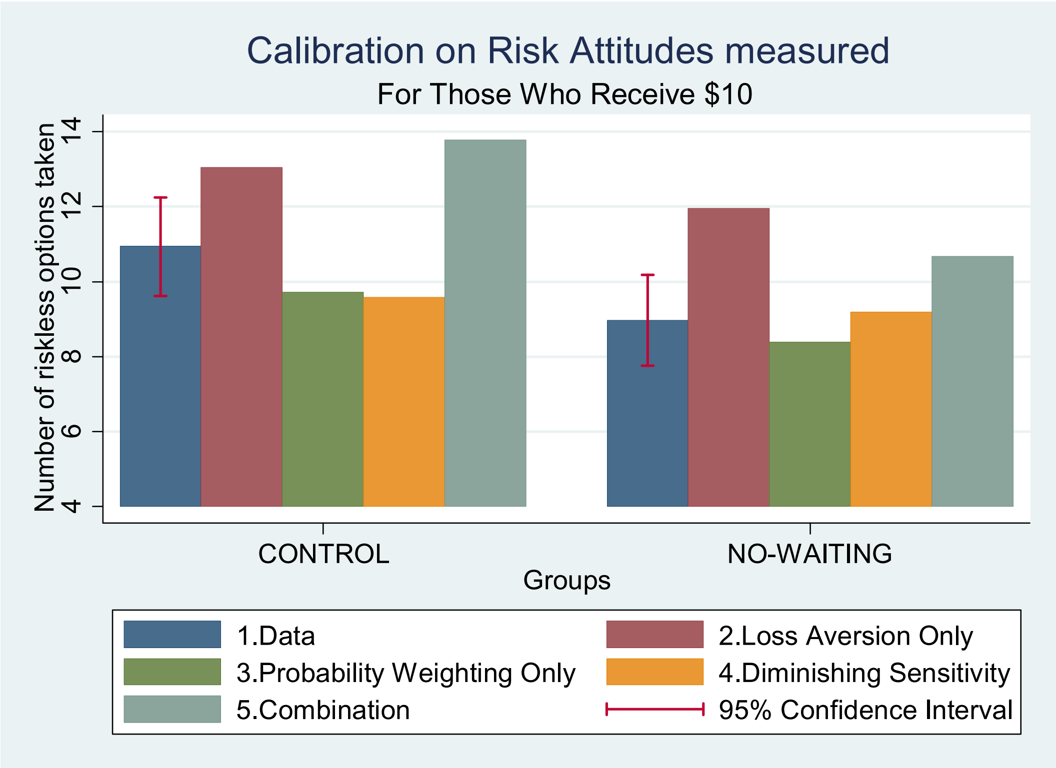



Note: The figure shows the risk attitudes for those who receive $20 in the control group and the treatment groups. The vertical axis stands for the number of riskless options taken in Holt and Laury table and it measures risk averse. The six bars in the control group and the treatment group stand for different measures from the tables. The detail is explained in the note from figure 2

Figure 5 Cumulative distributions of risk attitudes for those who receive $20




Note: The figure shows the cumulative distributions of risk attitudes for those who receive $20 in the three different groups. The horizontal axis stands for the average of all six measures from Holt and Laury table.




Note: I calibrate the risk attitudes for those who receive $10. I use the specification

$$u(x \mid RP) = \begin{cases} (x - RP)^{\alpha} & if \quad x \geq RP \\ -\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}$$

and the one-parameter form of probability weighting function: $\pi(p) = \exp(-(-\ln p)^{\gamma})$. The vertical axis represents the average number of riskless options taken from different measurements. In the control group, the reference points are $20 in the calibration. In the no-waiting group, the weights on stochastic reference points are 0.70 on $10, 0.15 on $15 and 0.15 on $20 in the calibration. There are five models in both the control group and the no-waiting group. The following is the list of models:

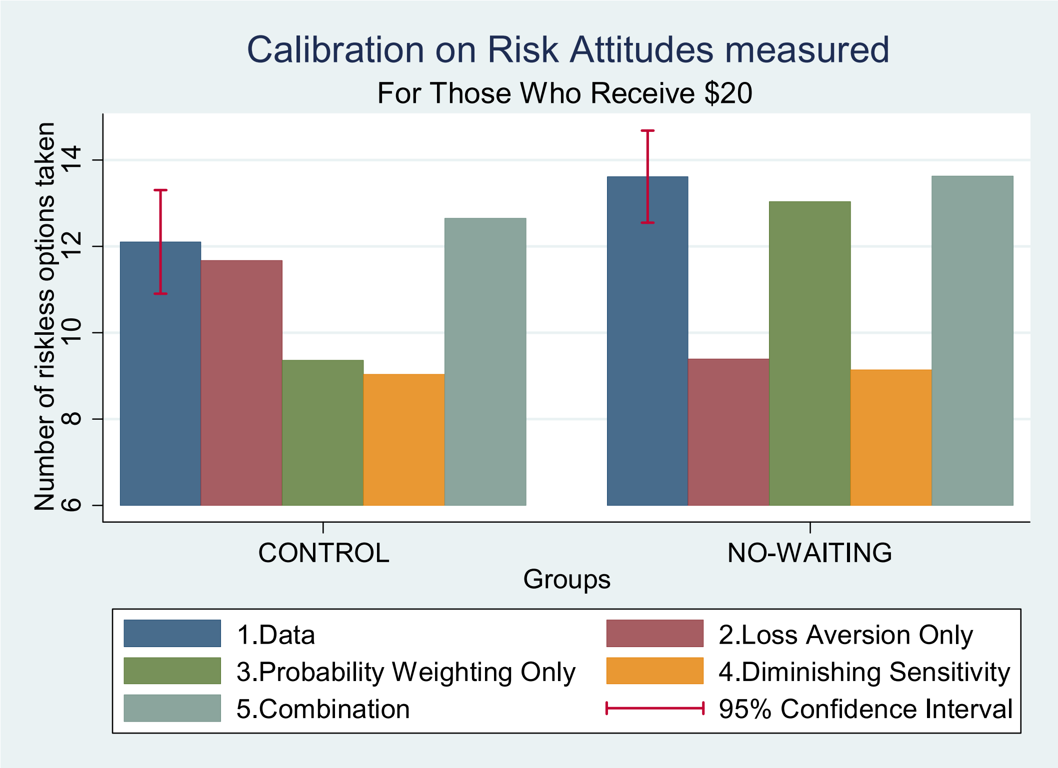
1st: Actual risk attitudes in my data

2nd: No diminishing sensitivity, loss aversion, linear probability weighting ($\alpha = \gamma = 1$, $\lambda = 1.81$)

3rd: No diminishing sensitivity, no loss aversion, nonlinear probability weighting ($\alpha = \lambda = 1$, $\gamma = 0.38$)

4th: Diminishing sensitivity, no loss aversion, linear probability weighting ($\lambda = \gamma = 1$, $\alpha = 0.92$)

5th: Diminishing sensitivity, loss aversion, nonlinear probability weighting ($\alpha = 0.92$, $\lambda = 1.81$ and $\gamma = 0.38$)




Note: I calibrate the risk attitudes for those who receive $20. I use the specification

$$u(x \mid RP) = \begin{cases} (x - RP)^{\alpha} & if \quad x \geq RP \\ -\lambda(RP - x)^{\alpha} & if \quad x < RP \end{cases}$$

and the one-parameter form of probability weighting function: $\pi(p) = \exp(-(-\ln p)^{\gamma})$. The vertical axis represents the average number of riskless options taken from different measurements. In the control group, the reference points are $20 in the calibration. In the no-waiting group, the weights on stochastic reference points are 0.70 on $10, 0.15 on $15 and 0.15 on $20 in the calibration. There are five models in both the control group and the no-waiting group. The following is the list of models:

1st: Actual risk attitudes in my data

2nd: No diminishing sensitivity, loss aversion, linear probability weighting ($\alpha = \gamma = 1, \lambda = 1.81$)

3rd: No diminishing sensitivity, no loss aversion, nonlinear probability weighting ($\alpha = \lambda = 1, \gamma = 0.38$)

4th: Diminishing sensitivity, no loss aversion, linear probability weighting ($\lambda = \gamma = 1, \alpha = 0.92$)

5th: Diminishing sensitivity, loss aversion, nonlinear probability weighting ($\alpha = 0.92$, $\lambda = 1.81$ and $\gamma = 0.38$)

**References:**


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Human (2011), "Reference Points and Effort Provision," *American Economic Review, 101(2):470-492.*

Agarwal, Sumit, John C Driscoll, Xavier Gabaix, and David Laibson (2011), "Learning in the Credit Card Market", *Available at SSRN: http://ssrn.com/abstract=1091623*

Ashraf, Nava, Dean Karlan, and Wesley Yin (2006), "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 121(2), pp. 635-672

Banerjee, Abhijit V. and Sendhil Mullainathan (2008), "Limited Attention and Income Distribution", *American Economic Review: Papers & Proceedings 2008, 98:2, 489–493*

Barr, Nicholas and Peter Diamond (2008), *Reforming Pensions: Principles and Policy Choices,* Oxford University Press, New York.

Barrientos, Armando, M.Ferreira, M.Gorman, A.Heslop, H.Legido-Quigley, P.Lloyd-Sherlock, V.Moller, S.Sabioa and M.Teixeira Vianna (2003) "Non-contributory pensions and poverty prevention. A comparative study of Brazil and South Africa," *IDPM/HelpAge International*.

Barseghyan, Levon, Molinari, Francesca, O'Donoghue, Ted and Teitelbaum, Joshua C. (2011), "The Nature of Risk Preferences: Evidence from Insurance Choices". Available at SSRN: http://ssrn.com/abstract=1646520

Banks, James, Cormac O'Dea and Zoe Oldfiel (2010), "Cognitive function, numeracy and retirement saving trajectories," *The Economic Journal, 120 (November), F381–F410.Doi: 10.1111/j.1468-0297.2010.02395.x.*

Bayer, Patrick J., B. Douglas Bernheim, and John Karl Scholz (2008), "The Effects of Financial Education in the Workplace: Evidence from a Survey of Employers," *Economic Inquiry*, 47 (4): 605–624.

Bell, David E., "Disappointment in Decision Making under Uncertainty," *Operations Research*, 1985, 33 (1), 1-27.

Bernheim, B. Douglas, (1998) "Financial Illiteracy, Education, and Retirement Saving," in Olivia S. Mitchell and Sylvester J. Schieber (eds.), *Living with Defined Contribution Pensions,* University of Pennsylvania Press, Pension Research Council, the Wharton School, University of Pennsylvania.

Beshears, John, James Choi, David Laibson, and Brigitte Madrian (2011), "Behavioral economics perspectives on public sector pension plans," *Journal of Pension Economics and Finance* 10(2): 315-336.

Cai, Jing (2011). "Social Networks and the Decision to Insure: Evidence from Randomized Experiments in China", *Working Paper, University of California at Berkeley*.

Carter, Michael R., Christopher B. Barrett, Stephen Boucher, Sommarat Chantarat, Francisco Galarza, John McPeak, Andrew Mude, and Carolina Trivelli (2008), "Insuring the Never Before Insured: Explaining Index Insurance Through Financial Education Games". Link:

http://i4.ucdavis.edu/files/briefs/2008-07-insuring-the-never-before-insured.pdf

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler (1997), "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics, 112 (2), 407-441.*

Carlin, B. I. and Robinson, D. T. (2011), "What does financial literacy training teach us?" *National Bureau of Economic Research Working Paper Series No. 16271*.

Carroll, Christopher, "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity, 1992, 1992 (2), 61-156,* and "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics, February 1997, 112 (1), 1-55.*

Choi, J., D. Laibson, B. Madrian, and A. Metrick. (2001), "Defined contribution pensions: Plan rules, Participant decisions, and the path of last resistance," *NBER Working Paper, No. W8655. Cambridge, Mass.: National Bureau of Economic Research.*

Choi, J., D. Laibson, B. Madrian, and A. Metrick. (2009), "Reinforcement Learning and Savings Behavior", *The Journal of Finance 64(6), December, 2515-2534.*

Cole, Shawn Allen, Gine Xavier, Tobacman, Jeremy Bruce, Topalova, Petia B., Townsend, Robert M. and Vickey, James I. (2011), "Barriers to Household Risk Management: Evidence from India," *Harvard Business School Finance Working Paper No. 09-116; FRB of New York Staff Report No.373. Available at SSRN: http://ssrn.com/abstract=1374076*

Cole, Shawn Allen, Thomas Sampson, and Bilal Zia (2010), "Money or Knowledge? What Drives demand for financial services in emerging markets?" *Journal of Finance, Forthcoming Harvard Business School Finance Working Paper No. 09-117*

Chamon, Marcos D. and Eswar S. Prasad (2010), "Why are Rates of Urban Households in China Rising?", *American Economic Journal: Macroeconomics 2010, 2:1, 93-130*

Crawford, Vincent P. and Juanjuan Meng (2011), "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review, 101(5): 1912–32.*

Deaton, Angus (1991), "Saving and Liquidity Constraints," *Econometrica, 59(5), 1221-1284.*

Deaton, Angus (1997), *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, The Johns Hopkins University Press: Baltimore.

Diamond, Peter, and Botond Koszegi (2003), "Quasi-hyperbolic discounting and retirement," *Journal of Public Economics 87 (9–10): 1839–72*.

Diamond, Peter, and Hausman, Jerry (1984), "Individual retirement and savings behavior," *Journal of Public Economics, vol. 23(1-2), pages 81-114*

Duflo, Esther (2000), "Child health and household resources in South Africa: Evidence from the old pension program," *The American Economic Review, 2000.*

Duflo, Esther, Michael Kremer and Jonathan Robinson (2009), "Nudging Farmers to Use Fertilizer: Evidence from Kenya," forthcoming, *American Economic Review*

Duflo, Esther and Emmanuel Saez (2003), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics (2003), Volume: 118, Issue: August, Publisher: National Bureau of Economic Research Cambridge, Mass., USA, Pages: 815-842.*

Edmonds, Eric V (2006), "Child labor and schooling responses to anticipated income in South Africa," *Journal of Development Economics 81(2006) 386-414*

Efron, B. and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*, Chapman & Hall: New York.

Eisenstein, Eric M. and Stephen J. Hoch (2005), "Intuitive Compounding: Framing, Temporal Perspective, and Expertise," (Under review at the *Journal of Consumer Research)*

Ericson, Keith M. Marzilli and Andreas Fuster (2009), "Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior," Working Paper.

Farber(2005), H. "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers," *Journal of Political Economy*, 113.

Farber, H. (2008) "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review, 98(2): 1069-1082*.

Fehr, Ernst and Lorenz Goette (2007), "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment", *the American Economic Review, Vol. 97, No. 1 (Mar., 2007), pp. 298-317*

Feng, Jin, He, Lixin and Sato, Hiroshi, Public Pension and Household Saving: Evidence from China (April 3, 2009). BOFIT Discussion Paper No. 2/2009. Available at SSRN: *http://ssrn.com/abstract=1379942*

Fredrickson, B. L., and Kahneman, D. (1993), "Duration neglect in retrospective evaluations of affective episodes," *Journal of Personality and Social Psychology, 65, 45–55.*

Gaurav, Sarthak, Shawn Cole, and Jeremy Tobacman (2011), "A randomized Evaluation of the impact of financial literacy on rainfall insurance take-up in Gujarat," *Research Paper No. 1.*

Gazzale, Robert, Julian Jamison, Alexander Karlan and Dean Karlan (2009), "Ambiguous Solicitation: Ambiguous Prescription", Williams College Economics Department Working Paper No. 2009-02. Available at SSRN: *http://ssrn.com/abstract=1371645*

Genesove, D. and C. Mayer (2001). "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics, 116(4), 1233-1260.*

Gerardi, Kristopher, Lorenz Goette and Stephan Meier (2010), "Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data," *Social Science Research Network*

Giné, Xavier, Robert Townsend and James Vickery (2008), "Patterns of Rainfall Insurance Participation in Rural India," *World Bank Economic Review, 22, 539-566.*

Ginè, X. and D. Yang (2009), “Insurance, credit, and technology adoption: Field experimental evidence from Malawi”, *Journal of Development Economics, 89, pp. 1-11*

Gourinchas, Pierre-Olivier and Jonathan Parker (2002), “Consumption over the Life Cycle,” *Econometrica, January 2002, 70 (1), 47-89*

Gul, Faruk (1991), “A Theory of Disappointment Aversion," *Econometrica*, 59 (3), 667-686.

Guo, Ping and Gang Chen (2006), *Data Analysis of Sampling Survey of the Aged Population in Urban/Rural China 2006,* Zhongguo she hui Press, Beijing

Haselhuhn, Michael P., Devin G. Pope, Maurice E. Schweitzer (2009), “Size Matters (and so Does Experience): How Personal Experience with a Fine Influences Behavior”, *Available at SSRN: http://ssrn.com/abstract=1270746*

Heffetz, Ori and John List (2012) “Assignment, Expectations, and Endowment: Variations in the Lab and Their Effects on Choice,” *Working Paper*

Holt, Charles A. and Susan K. Laury (2002), “Risk Aversion and Incentive Effects," *The American Economic Review, 92 (5), 1644-1655.*

Holt, Charles A. Markets (2005), “Games, and Strategic Behavior: Recipes for Interactive Learning.” *Unpublished Manuscript*

Jensen, Robert (2000), “Agricultural Volatility and Investments in Children”, *the American Economic Review; May 2000; 90, 2; ABI/INFORM Global, pg 399.*

Kahneman, Daniel and Amos Tversky (1979), “Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, *47 (2), 263-291.*

Kling, Jeffery R., Jeffery B. Liebman, and Lawrence F. Katz (2007), “Experimental Analysis of Neighborhood effects,” *Econometrica, Vol. 75, No. 1 (January, 2007), 83–119*

Koszegi, Botond and Matthew Rabin (2006), “A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, *121 (4), 1133-1165*.

- And (2007) “Reference-Dependent Risk Attitudes," *The American Economic Review*, *97 (4), 1047-1073.*
- And (2009), “Reference-Dependent Consumption Plans," *American Economic Review*, *2009, 99 (3), 909-936.*

Laibson, David I., Andrea Repetto, and Jeremy Tobacman (1998), “Self-Control and Saving for Retirement,” Brookings Papers on Economic Activity, 91–196

Levitt, Steven D. and John A List (2007), “What Do Laboratory Experiments Tell Us About the Real World?” *Journal of Economic Perspectives* 21(2):153-74.

Liebman, J. and R. Zeckhauser (2008), “Simple Humans, Complex Insurance, Subtle Subsidies,” *NBER Working Paper No. 14330*

Loomes, Graham and Robert Sugden (1986), “Disappointment and Dynamic Consistency in Choice under Uncertainty," *Review of Economic Studies*, *53 (2), 271-82.*

Lusardi, Annamaria (1999), "Information, Expectations, and Savings for Retirement," in Henry Aaron (ed.), *Behavioral Dimensions of Retirement Economics*, Washington, D.C.: Brookings Institution Press and Russell Sage Foundation, 1999, pp. 81-115.

Lusardi, Annamaria, and Olivia S. Mitchell (2007a), "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 54:1, pp.205-24.

Lusardi, Annamaria, and Olivia S. Mitchell (2007b), "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education," *Business Economics*, 42(1): 35-44.

Lusardi, A., and P. Tufano (2008), "Debt Literacy, Financial Experiences and Overindebtedness," *Working Paper 14808, National Bureau of Economic Research (NBER), Cambridge, MA.*

Lybbert, T.J., C.B. Barrett, S. Boucher, M.R. Carter, S. Chantarat, F. Galarza, J. G. McPeak and A.G. Mude. (2009), "Dynamic Field Experiments in Development Economics: Risk Valuation in Kenya, Morocco and Peru," *Agricultural and Resource Economics Review* 39(2): 176-192 (2010)

Madrian, B., and D. Shea (2001), "The power of suggestion: Inertia in 401(k) participation and savings behavior," *Quarterly Journal of Economics 116(4): 1149-87.*

Malmendier, U. and Nagel, S. (2011), "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" Quarterly Journal of Economics*, 126 (2011), 373–416.*

Malmendier, U. and Nagel, S. (2011), "Learning from Inflation Experiences", *Working paper*

McKenzie, C. R. M., and Liersch, M. J. (2011), "Misunderstanding savings growth: Implications for retirement savings behavior," *Journal of Marketing Research*

Meier, Stephan, and Charles Sprenger (2008), "Discounting Financial Literacy: Time Preferences and Participation in Financial Education Programs," Unpublished

National Bureau of Statistics of P.R. China, *Chinese Statistical Yearbook*, Beijing: China Statistics Press; 2000–2010.

O'Donoghue, Ted, and Matthew Rabin (1999), "Procrastination in preparing for retirement," In *Behavioral dimensions of retirement economics*, ed. Henry Aaron, 125–56, Washington DC: The Brookings Institute.

Odean, Terrance(1998), "Are Investors Reluctant to Realize Their Losses?" *The Journal of Finance*, *53 (5), 177-1798.*

Orszag, Peter R. and Joseph E. Stiglitz (2001), "Rethinking Pension Reform: Ten Myths About Social Security Systems," in *New Ideas about Old Age Security*, Robert Holtzman and Joseph E. Stiglitz, eds., Washington, DC: World Bank, pp. 17-56.

Plott, Charles R. (2001), Public Economics, Political Processes and Policy Applications. Collected Papers on the Experimental Foundations of Economics and Political Science. Volume One. *Cheltenham, UK: Edward Elgar Publishing.*

Pope, Devin and Maurice Schweitzer (2011), "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *American Economic Review*, *101(1):129-57*

Post, Thierry, Martihn J. van den Assem, Guido Baltussen, and Richard H.Thaler (2008), "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game

Show," *American Economic Review*, *98 (1), 38-71.*

Prelec, Drazen (1998), "The Probability Weighting Function," *Econometrica*, *66 (3), 497-527*

Rosenzweig, Mark R. and Wolpin, Kenneth I. (1993), "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-income Countries: Investments in Bullocks in India", *the Journal of Political Economy, Vol. 101, No. 2 (Apr., 1993), pp. 223-244*

Samuelson, William and Richard Zeckhauser, "Status Quo Bias in Decision Making," Journal of Risk and Uncertainty, 1 (1988), 7-59.

Samwick, Andrew A. (1998), "Discount Rate Heterogeneity and Social Security Reform," *Journal of Development Economics, 57 (1), 117-146.*

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun (2006), "Are Americans Saving Optimally for Retirement?" *Journal of Political Economy 114 (August): 607–43*.

Schreiber, C. A., and Kahneman, D. (2000), "Determinants of the remembered utility of aversive sounds," *Journal of Experimental Psychology: General, 129, 27– 42.*

Schwartzer, H. and A. Querino (2002), "Non-contributory pensions in Brazil: the impact on poverty reduction," *ESS Paper 11, ILO Social Security Policy and Development Branch.*

Shi, S.-J. (2006), "Left to market and family – again? Ideas and the development of the rural pension policy in China", *Social Policy and Administration, Vol. 40 No. 7, pp. 791-806*.

Song, Jian (2001), "A Comprehensive Review of Studies on Old - supporting Problems in Rural China," *Population Research, 25 (6)*

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Point", UCSD

Stango, Victor, and Jonathan Zinmam (2009), "Exponential Growth Bias and Household Finance," *The Journal of Finance*, *Volume 64, Issue 6, pages 2807–2849, December 2009*

State Council of the People's Republic of China (2006) *The Development of China's Undertakings for the Aged*, http://www.china.org.cn/english/aged/192020.htm

Sun Yefang Economic Science Fund Association (2010), "Survey on Suicides of the Rural Elderly in China," http://www.21bcr.com/a/zhuan__ti/laonian/2010/0606/672.html

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen (2010), "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review, 100(1): 557–71*

Thaler, Richard H., and Cass R. Sunstein. (2003), "Libertarian Paternalism." *American Economic Review, Papers and Proceedings* 93, no. 2 (May): 175–79.

———(2008), *Nudge: Improving Decisions about Health, Wealth, and Happiness,* Yale University Press.

Tversky, Amos and Craig R. Fox(1995), "Weighing Risk and Uncertainty," *Psychological Review*, *102 (2), 269-283.*

—and Daniel Kahneman(1992), "Advances in Prospect Theory: Cumulative

Representation of Uncertainty," Journal of Risk and Uncertainty, *5 (4), 297-323.*

United Nations (2009), "Population Aging and Development 2009," United Nations, New York, http://www.un.org/esa/population/publications/ageing/ageing2009.htm

United Nations (2011), "World Population Prospects, The 2011 Revision," United Nations, New York, http://esa.un.org/unpd/wpp/index.htm

Venti, Steven and David A. Wise (1996), "The Wealth of Cohorts: Retirement Saving and the Changing Assets of Older Americans," *NBER Working Papers 5609*

Wang, Dewen (2006), "China's Urban and Rural Old Age Security System: Challenges and Options," *China & World Economy 102 – 116, Vol. 14, No. 1*

Wang, Mei and Xia, Chuanling (1994), "Analysis of Beijing middle-aged family pension," *Population Research, 19(4)*

Wang, Yicai (2000), "Family pension, land security and social insurance is a combination of a natural choice to solve the rural old-age", *Population Research, 25(5)*

Municipal Bureau of Statistics (2011), *Data for Districts and Counties in Shaanxi*.

Zhang,Wenjuan, and Tang, Dan (2008), "The New Rural Social Pension Insurance Programme of Baoji City," HelpAge International-Asia/Pacific, Chiang Mai.

Zhao, Yaohui, John Strauss, Albert Park and Yan Sun (2009), "China Health and Retirement Longitudinal Study, Pilot, User's Guide," *National School of Development, Peking University*

## Appendix A: Numerical Solution of Consumer Problem

### A.1 Timing of event in the life-cycle model

An individual has cash on hand $X_t$ at the beginning of age *t*, and consumes $C_t$ during age *t*. At the end of age *t*, the remaining cash on hand is $X_t - C_t - Q_t$. At the beginning of age *t+1*, we first resolve the lifetime uncertainty and then resolve income uncertainty, if individual survives. Nature takes a draw with probability $p_t$ that the individual survives in age *t+1*. If individual survives, nature takes a draw of income $Y_{t+1}$ according to the income process. The individual also receives the return from assets, $R(X_t - C_t - Q_t)$ and pension $Z_{t+1}$ if they are age 60 or over. Therefore, the individual has cash on hand $X_{t+1}$ at the beginning of age *t+1*.

### A.2 Estimation of Exogenous Process

Survival probabilities are based on 2009 life tables from the World Health Organization ( http://www.who.int/healthinfo/statistics/mortality_life_tables/en/). Survival probabilities can be calculated at any age by simply dividing the number of survivors at the terminal age by the number at the beginning age. The data has only five year intervals, so we interpolate the survival probability at each age using Piecewise cubic Hermite interpolation to preserve the shape of the data.

Income uncertainty and age-specific income growth are estimated from the China Health and Nutrition Study (CHNS) (http://www.cpc.unc.edu/projects/china), a large scale longitudinal survey conducted in nine provinces of China in 1989, 1991, 1993, 1997, 2000, 2004, 2006 and 2009. The survey covers coastal, middle, northeastern, and western provinces. The CHNS also includes cities with different income levels, and surveys both rural and urban residents. Although the CHNS is not a nationally representative sample, the provinces covered vary substantially in terms of geography and economic development. The CHNS collects information on a wide range of individual socioeconomic, health and nutritional characteristics. The CHNS also includes information on income and wealth, which is the key information we use in our study.

We estimate the variance of the permanent and transitory component of shocks, $\sigma_n^2$ and $\sigma_u^2$, using CHNS and the methodology of Carroll and Samwick (1997).

To estimate the age-specific expected income growth, we need to decompose age, cohort, and year effects from the panel data, and to construct age-profiles of income. As discussed in Deaton (1997), it is not possible to decompose these three effects without further restrictions. This follows from the identity that interview year less age equals birth year. Following Deaton (1997), we define year dummies in a way that makes the year effects orthogonal to a time trend:

$$D_t = d_t - [(t-1)d_2 - (t-1)d_1]$$

where *t=3,…,T*, $d_t$ is the usual year dummy, equal to 1 if the year is t and 0 other wise.

We then estimate the following regression:

$$\ln Y_i = a_i \pi_1 + a_i^2 \pi_2 + c_i \pi_3 + D_i \pi_4 + f_i \pi_5 + u$$

where $a_i$ is the age, $c_i$ is a complete set of cohort dummies (less the middle one), and $f_i$

is the household size. The coefficients of the regressions give the third through final year coefficients; the first and second can be recovered from the fact that all year effects add up to zero.

With these estimates, we construct household-level income uncontaminated by cohort and time effects:

$$\ln \hat{Y}_i = a_i \hat{\pi}_1 + a_i^2 \hat{\pi}_2 + \bar{f}_i \hat{\pi}_5 + \hat{u}$$

$\ln \hat{Y}_i$ is the income of household with family size $\bar{f}_i$ and born in the middle cohort. The average age-profiles of income can be constructed by averaging these data across households:

$$\ln \hat{Y}_a = a \hat{\pi}_1 + a^2 \hat{\pi}_2 + \bar{f} \hat{\pi}_5$$

We can calculate the expected income growth rate by first differencing the log-average income.

### A.3 Consumption Rules

We solve the optimal consumption rule by solving the Euler equation. We start at age *N*, assumed to be 100, and solve the Euler equation with all possible states (the problem at this stage is trivial, since the household will simply consume all income). We move backward to the previous period and solve for the consumption rule by the Euler equation. We go all the way to the starting age S and consequently recover the age-specific consumption rules.

The problem consists in evaluating the expectation. Since *N* and *U* are log normally distributed, a natural way to evaluate these integrals is to perform a two dimensional Gauss-Hermite quadrature using the product rule:

$$\begin{aligned} \mathrm{E}[u'(c_{t+1}(x_{t+1})G_{t+1}N_{t+1})] &= \int u'(c_{t+1}(x_{t+1})G_{t+1}N_{t+1})dF(N)dF(U) \\ &= \int_{-\infty}^{\infty} f_t(n,u)e^{-n^2}e^{-u^2}dndu \\ &\approx \sum_{i,j} \omega_i \omega_j f_t(n_i,u_j) \end{aligned}$$

where $f_t(n,u) = \frac{1}{\pi}u'(c_{t+1}((x_t - c_t)\frac{R}{G_{t+1}}e^{-\sqrt{2}\sigma_n n} + e^{\sqrt{2}\sigma_u u})G_{t+1}e^{\sqrt{2}\sigma_u u}$ , $u = \frac{\log(U)}{\sqrt{2}\sigma_u}$ and $n = \frac{\log(N)}{\sqrt{2}\sigma_n}$.

The weights $\omega_{ij}$ and nodes $n_i$, $u_j$ are tabulated in Judd (1998). In practice, we performed a quadrature with 10 nodes.

We use a standard discretization method to solve the optimal consumption rule. We specify an exogenous grid for cash on hand, $\{x^j\}_{j=1}^{J} \subset [0, x^{\max}]$. In practice, for each value on the grid, $x^j$, we find the associated consumption, $c^j$, that satisfies the Euler equation. We constrain the associated consumption to be positive and less than *x*. In practice, with 50 points on the grid and 80 time periods, we must solve the Euler equation 4,000 times. Consumption will be evaluated using interpolation or extrapolation methods.

Then we simulate optimal consumption (and therefore wealth) each period for each household by simulating income. Consider a household *h* with age *S*, the first working age. The household is assumed to begin with zero assets and zero income. We then simulate the income according to the income process in Equation (3), and calculate consumption in age *S* according to the consumption rule in age S. We move forward to the next period to simulate income and calculate consumption until we have a complete consumption path. For those with age $t > S$, their initial assets are assumed to be the wealth of household *h* at age *t*. Thus, we can simulate optimal consumption path for each household.

### A.4 Bootstrapping the Confidence Interval

We bootstrap the confidence interval of the calibrated contribution levels. The detailed block bootstrapping procedure takes the following steps:

1. Choose the block. We assume each village is independent, and choose the village as a block.
2. Resample the blocks and generate a bootstrap resample. The number of villages in the bootstrap resample is the same as in the original data.
3. Calibrate the contribution level for the bootstrap resample. Given the optimal life-cycle consumption path, for the households in the resample we can calculate the lifetime utility for each choice of contribution level, and thus find the optimal contribution level in the rural pension program.
4. We resample the blocks for NB=100 and calculate the mean and confidence interval.

### Appendix B: Measure of Financial Literacy

We use absolute distance to the correct answer to measure how close the respondents' answers are to the correct ones. The absolute distance for each individual and each question is calculated in the following formula:

$$E(|x - x_c|) = \int_{x_l}^{x_u} |x - x_c| f(x) dx$$

where $x$ is the chosen answer and $x_c$ is the correct answer. Since all the questions are multiple-choice questions, we assume the chosen answer $x$ is a uniform distribution on $[x_l, x_u]$, where $x_l$ and $x_u$ are two boundaries of the chosen option. For example, the correct answer for Question 3 is $x_c = 574$. If a subject choose the option 1 (100-300 RMB), we assume his/her answer is a uniform distribution on [100, 300]. Then we apply the Monte Carlo integration method to calculate the expected absolute distance. Note that the options for Questions 4 and 5 are qualitative, and thus we just measure whether a subject answers them correctly. If a subject does not answer the question, we assume the absolute distance is the same as that of the worst option in the question in order to distinguish him/her from those who answer the question. We reverse the sign for absolute distance so that a higher value of the measure represents an answer closer to the correct one.

## Appendix C: Simulation and Estimation Procedures

### C.1 Simulation of Insurance Take-up under Standard Model

We simulate the take-up decisions with the following steps:

1. Take a uniform draw of $\alpha$ from the interval according to each household's choices of riskless options
2. Take two extreme type I error term and difference them to get logistic error term
3. Use the draw of $\alpha$ ,self-reported p and the error term to calculate the insurance decision of each household and the percentage of take-up in the simulated sample
4. Repeat 1 to 3 for 100 times and calculate the mean and standard deviation of take-up.

**C.2 MLE Estimation of Weight Parameters**

We estimate $\mu_1$ and $\mu_2$ with MLE and simulation with the following steps:

1. Take a uniform draw of $\alpha$ from the interval according to each household's choices of riskless options
2. Constrain $\alpha$ to be the draw value and $p$ to be the perceived probability of future disasters from our survey data, then estimate $\mu_1$ and $\mu_2$ with MLE
3. Redo step 1 and 2 for 100 times to generate 100 $\mu_1$ and $\mu_2$
4. Compare the distribution of $\mu_1$ and $\mu_2$

**Table A1. Population Aging in China**

| | 1970 | 1990 | 2010 | 2030 | 2050 |
|---|---|---|---|---|---|
| Population (thousands) | 814,623 | 1,145,195 | 1,341,335 | 1,393,076 | 1,295,604 |
| Population growth rate (%) | 2.74 | 1.61 | 0.51 | -0.03 | -0.55 |
| Life expectancy at birth | 59.4 | 68.9 | 72.7 | 76.4 | 79.1 |
| Percentage aged 60 or over | 6.6 | 8.9 | 12.3 | 24.4 | 33.9 |

Source: United Nations

*The number is the average of five years before the year

**Table A2. Risk Attitude and Time Preference Qustions**

| | Option A | Option B |
|---|---|---|
| **Panel A: Risk Attitude** | | |
| 1 | 50 RMB | Toss a coin. If it is heads, you get 200RMB. If it is tails, you get nothing. |
| 2 | 80 RMB | Toss a coin. If it is heads, you get 200RMB. If it is tails, you get nothing. |
| 3 | 100RMB | Toss a coin. If it is heads, you get 200RMB. If it is tails, you get nothing. |
| 4 | 120RMB | Toss a coin. If it is heads, you get 200RMB. If it is tails, you get nothing. |
| 5 | 150RMB | Toss a coin. If it is heads, you get 200RMB. If it is tails, you get nothing. |
| **Panel B: Time Preference** | | |
| 1 | 1000 RMB today | 1063 RMB in one year |
| 2 | 1000 RMB today | 1188 RMB in one year |
| 3 | 1000 RMB today | 1313 RMB in one year |
| 4 | 1000 RMB today | 1437 RMB in one year |
| 5 | 1000 RMB today | 1563 RMB in one year |
| 6 | 1000 RMB today | 1688 RMB in one year |
| 7 | 1000 RMB in 2 years | 1063 RMB in 3 years |
| 8 | 1000 RMB in 2 years | 1188 RMB in 3 years |
| 9 | 1000 RMB in 2 years | 1313 RMB in 3 years |
| 10 | 1000 RMB in 2 years | 1437 RMB in 3 years |
| 11 | 1000 RMB in 2 years | 1563 RMB in 3 years |
| 12 | 1000 RMB in 2 years | 1688 RMB in 3 years |

Note: Risk attitudes were elicited for all the households with questions in Panel A. For those who were assigned to the Education group, risk attitudes were elicited after the education. Households were asked to make five hypothetical decisions to choose between riskless option A and risky option B. We use the number of riskless options as a measurement of risk averse. The more the riskless options are chosen, the more the risk averse is. Time preferences were elicited for all the households with questions in Panel B. For those who were assigned to the Education group, time preferences were elicited after the education. Time preferences were elicited by asking sample households to choose between receiving some amount of money now (option A) and increasing amount of money one year later (option B) in Table A2. We use the number of patient options (option B) as a measurement of patience. The more the patient options are chosen, the more the patience it is

**Table A3. Range of Risk Aversion and Time Prference**

**Panel A: Risk Attitude**

| Number of riskless options | Number of observation | Range of α for CRRA $u(x)=x^{1-\rho}/(1-\rho)$ |
|---|---|---|
| 0 | 117 | $\rho<-1.4$ |
| 1 | 19 | $-1.4<\rho<-0.35$ |
| 2 | 57 | $-0.35<\rho<0$ |
| 3 | 83 | $0<\rho<0.25$ |
| 4 | 61 | $0.25<\rho<0.5$ |
| 5 | 763 | $\rho>0.5$ |

**Panel B: Time Preference**

| Number of patient options | Number of observation | Range β |
|---|---|---|
| 0 | 442 | $\beta<0.59$ |
| 1 | 25 | $0.59<\beta<0.64$ |
| 2 | 59 | $0.64<\beta<0.70$ |
| 3 | 79 | $0.70<\beta<0.76$ |
| 4 | 83 | $0.76<\beta<0.84$ |
| 5 | 75 | $0.84<\beta<0.94$ |
| 6 | 334 | $\beta>0.94$ |

**Table A.1 The Paired Lottery-Choice Decisions with Probability Changing**

| Option 1 | | Option 2 | | | Expected payoff |
|---|---|---|---|---|---|
| Payoff | Probability | Payoff | Probability | Payoff | difference |
| 10 | 10% | 15 | 90% | 5 | 4 |
| 10 | 15% | 15 | 85% | 5 | 3.5 |
| 10 | 20% | 15 | 80% | 5 | 3 |
| 10 | 25% | 15 | 75% | 5 | 2.5 |
| 10 | 30% | 15 | 70% | 5 | 2 |
| 10 | 35% | 15 | 65% | 5 | 1.5 |
| 10 | 40% | 15 | 60% | 5 | 1 |
| 10 | 45% | 15 | 55% | 5 | 0.5 |
| 10 | 50% | 15 | 50% | 5 | 0 |
| 10 | 55% | 15 | 45% | 5 | -0.5 |
| 10 | 60% | 15 | 40% | 5 | -1 |
| 10 | 65% | 15 | 35% | 5 | -1.5 |
| 10 | 70% | 15 | 30% | 5 | -2 |
| 10 | 75% | 15 | 25% | 5 | -2.5 |
| 10 | 80% | 15 | 20% | 5 | -3 |
| 10 | 85% | 15 | 15% | 5 | -3.5 |
| 10 | 90% | 15 | 10% | 5 | -4 |
| 10 | 95% | 15 | 5% | 5 | -4.5 |
| 10 | 100% | 15 | 0% | 5 | -5 |

Note: In Table A1 I fix the payoffs but change the probabilities in risky options. For each question in a row, subjects are asked to choose between option 1 and option 2. Subjects cannot see the expected payoff difference.

**Table A.2 The Paired Lottery-Choice Decisions with Payoff Changing**

| Option 1 | | Option 2 | | | Expected payoff |
| --- | --- | --- | --- | --- | --- |
| Payoff | Probability | Payoff | Probability | Payoff | difference |
| 10 | 50% | 10 | 50% | 5 | 2.5 |
| 10 | 50% | 10.5 | 50% | 5 | 2.25 |
| 10 | 50% | 11 | 50% | 5 | 2 |
| 10 | 50% | 11.5 | 50% | 5 | 1.75 |
| 10 | 50% | 12 | 50% | 5 | 1.5 |
| 10 | 50% | 12.5 | 50% | 5 | 1.25 |
| 10 | 50% | 13 | 50% | 5 | 1 |
| 10 | 50% | 13.5 | 50% | 5 | 0.75 |
| 10 | 50% | 14 | 50% | 5 | 0.5 |
| 10 | 50% | 14.5 | 50% | 5 | 0.25 |
| 10 | 50% | 15 | 50% | 5 | 0 |
| 10 | 50% | 15.5 | 50% | 5 | -0.25 |
| 10 | 50% | 16 | 50% | 5 | -0.5 |
| 10 | 50% | 16.5 | 50% | 5 | -0.75 |
| 10 | 50% | 17 | 50% | 5 | -1 |
| 10 | 50% | 17.5 | 50% | 5 | -1.25 |
| 10 | 50% | 18 | 50% | 5 | -1.5 |
| 10 | 50% | 18.5 | 50% | 5 | -1.75 |
| 10 | 50% | 19 | 50% | 5 | -2 |

Note: In Table A2 I fix the probabilities but change the payoffs in risky options. For each question in a row, subjects are asked to choose between option 1 and option 2. Subjects cannot see the expected payoff difference.

**Table A.3 The Paired Lottery-Choice Decisions with Payoff Changing**

| Option 1 | | Option 2 | | | Expected payoff |
|---|---|---|---|---|---|
| Payoff | Probability | Payoff | Probability | Payoff | difference |
| 10 | 50% | 20 | 50% | 10 | -5 |
| 10.5 | 50% | 20 | 50% | 10 | -4.5 |
| 11 | 50% | 20 | 50% | 10 | -4 |
| 11.5 | 50% | 20 | 50% | 10 | -3.5 |
| 12 | 50% | 20 | 50% | 10 | -3 |
| 12.5 | 50% | 20 | 50% | 10 | -2.5 |
| 13 | 50% | 20 | 50% | 10 | -2 |
| 13.5 | 50% | 20 | 50% | 10 | -1.5 |
| 14 | 50% | 20 | 50% | 10 | -1 |
| 14.5 | 50% | 20 | 50% | 10 | -0.5 |
| 15 | 50% | 20 | 50% | 10 | 0 |
| 15.5 | 50% | 20 | 50% | 10 | 0.5 |
| 16 | 50% | 20 | 50% | 10 | 1 |
| 16.5 | 50% | 20 | 50% | 10 | 1.5 |
| 17 | 50% | 20 | 50% | 10 | 2 |
| 17.5 | 50% | 20 | 50% | 10 | 2.5 |
| 18 | 50% | 20 | 50% | 10 | 3 |
| 18.5 | 50% | 20 | 50% | 10 | 3.5 |
| 19 | 50% | 20 | 50% | 10 | 4 |

Note: In Table A3 I change the payoffs in riskless options. For each question in a row, subjects are asked to choose between option 1 and option 2. Subjects cannot see the expected payoff difference.

**Table A.4 Summary Questions for Each Table**

| | Summary Questions |
|---|---|
| After Table A1 | Now you have a choice between (1) Keep the $10 (2) Take the following bet: p% probability to get $15 and (100-p) % probability to get $5. What is the minimum probability p% that you will choose choice option 2? |
| After Table A2 | Now you have a choice between (1) Keep the $10 (2) Take the following bet: 50% probability to get $X and 50% probability to get $5. What is the minimum X that you will choose choice 2? |
| After Table A3 | Now you have a choice between (1) Keep $X (2) Take the following bet: 50% probability to get $10 and 50% probability to get $20. What is the minimum X that you will choose choice 1? |