UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Three Essays on Insurance Demand and Impact

Permalink
https://escholarship.org/uc/item/98h0n85t

Author
Cai, Jing

Publication Date
2012

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Three Essays on Insurance Demand and Impact

by

Jing Cai

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alain de Janvry, Co-Chair
Professor Elisabeth Sadoulet, Co-Chair
Professor Edward Miguel
Professor David Levine

Spring 2012

Three Essays on Insurance Demand and Impact

Copyright 2012
by
Jing Cai

Abstract

Three Essays on Insurance Demand and Impact

by

Jing Cai

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Alain de Janvry, Co-Chair
Professor Elisabeth Sadoulet, Co-Chair

Some new technologies or financial products have the potential to dramatically improve economic development and household welfare, but adoption is often sub-optimally slow. In this dissertation, I explore the barriers to the diffusion of innovations, identify and assess potential ways to overcome these constraints, and evaluate the impact of innovation adoption on household behavior, using both experimental and non-experimental methods.

Using data from a two-year randomized experiment in rural China, the first chapter studies the influence of social networks on the decision to adopt a new weather insurance product and the mechanisms through which social networks operate. In the first year, I provided financial education to a random subset of farmers and found a large social network effect on insurance take-up: for untreated farmers, having an additional friend receiving financial education raises take-up by almost half as much as obtaining financial education directly, a spillover effect equivalent to offering a 12% reduction in the average insurance premium. By varying the information available to subjects about their peers' take-up decisions and using randomized default options, I show that the positive social network effect is not driven by scale effects, imitation, or informal risk-sharing, but instead by the diffusion of insurance knowledge. One year later, social networks continue to affect insurance demand: observing an above-median share of friends receiving payouts increases insurance take-up at a rate equivalent to about 50% of the impact of receiving payouts directly. I also find that social network effects are larger in villages where households are more strongly connected, and when the people who receive financial education first are more central in the social network.

The second chapter is based on a coauthored paper with Changcheng Song, "Insurance Take-up in Rural China: Learning from Hypothetical Experience". This chapter uses a novel experimental design to test for the role of experience and information in insurance take-up in rural China, where weather insurance was a new and highly subsidized product. We randomly select a group of poor households to play insurance games and find that it improves the actual insurance take-up by 48%. In order to determine the mechanism behind

this effect, we test whether it 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) hypothetical experience of disaster. We show that the effect cannot be explained by mechanisms (1) to (3), and that the experience acquired in playing the insurance game matters. We develop a simple model in which agents give less weight to disasters and benefits which they experienced infrequently. Our estimation also suggests that compared with experience with real disasters in the previous year, experience gained in the insurance game played recently has a stronger effect on the actual insurance take-up, implying that learning from experience displays a strong recency effect.

In the third chapter, I take advantage of a natural experiment and a rich household-level panel dataset to test the impact of an agricultural insurance program on household level production, borrowing, and saving. The empirical strategy includes both difference-in-difference and triple difference estimations. I find that first, introducing insurance increases the production area of insured crops by around 20%, and it decreases production diversification; second, provision of insurance raises the credit demand by 25%; third, it decreases the household saving by more than 30%; fourth, the effect of insurance policy on borrowing persists in the long-run, while that on saving is significant only in the medium-run; fifth, the impact of insurance is bigger on larger farmers, and households with lower migration remittance.

To my parents

Contents

1 Social Networks and the Decision to Insure **1**
1.1 Introduction . . . 1
1.2 Background . . . 5
1.3 Experimental Design and Data . . . 6
1.3.1 Year One: Identify the Short-run Social Network Effect . . . 6
1.3.2 Year Two: Identify the Social Network Effect Over Time . . . 8
1.3.3 Data, Summary Statistics and Randomization Check . . . 9
1.4 Estimation Strategies and Results . . . 11
1.4.1 Effect of Social Networks on Initial Year Insurance Demand . . . 11
1.4.2 Effect of Social Networks on Second Year Insurance Demand . . . 21
1.5 Conclusions . . . 25

2 Insurance Take-up and Learning from Experience **46**
2.1 Introduction . . . 46
2.2 Rice Insurance in China . . . 49
2.3 Experimental Design and Survey Data . . . 50
2.3.1 Experimental Design . . . 50
2.3.2 Survey Data . . . 52
2.4 Empirical Result . . . 53
2.4.1 The Impact of Hypothetical Experience on Actual Take-up . . . 53
2.4.2 Possible Channels . . . 54
2.5 Models . . . 58
2.5.1 Standard Model . . . 58
2.5.2 Model Based on Experience . . . 60
2.6 Conclusions . . . 61

3 The Impact of Insurance Provision **74**
3.1 Introduction . . . 74
3.2 Background . . . 76
3.3 Theoretical Model . . . 77
3.3.1 Two-period model when insurance is not provided . . . 78

3.3.2 Two-period model when insurance is provided . . . 79
3.3.3 Combine the two models . . . 80
3.3.4 Break-even conditions of the bank . . . 80
3.3.5 Conclusion of the model . . . 81
3.4 Data and Summary Statistics . . . 82
3.5 Estimation Strategies and Results . . . 85
3.5.1 Empirical Strategies . . . 85
3.5.2 Estimation Results . . . 87
3.6 Conclusions . . . 88

A Social Networks and the Decision to Insure **101**
A.1 Supplementary Figures and Tables . . . 101
A.2 Glossary of Network Terminology . . . 108
A.3 An Insurance Demand Model . . . 109
A.3.1 Individual Insurance Demand . . . 109
A.3.2 Aggregate Insurance Demand . . . 110

B Insurance Take-up in Rural China **115**
B.1 Simulation Methods . . . 115
B.1.1 Simulation of Insurance Take-up under Standard Model . . . 115
B.1.2 MLE Estimation of Weight Parameters . . . 115

C The Impact of Insurance Provision **119**
C.1 Two-period model when insurance is not provided . . . 119
C.2 Two-period model when insurance is provided . . . 120

Acknowledgments

The completion of this dissertation would not have been possible without the support and guidance provided by my two advisors, Alain de Janvry and Elisabeth Sadoulet. I cannot thank Alain and Betty enough for their invaluable mentoring over the last six years, which has given me tools that I will continue to use and value. I am grateful to David Levine and Edward Miguel for their advice, guidance, and relentless encouragement. David and Ted have helped in so many ways to shape this work while continue to provide firm support and constructive feedback along the way. I am also indebted to Ann Harrison for her help and advice, not only on my research but also on my career development. Ann has been a tremendous help during my time at Berkeley and the World Bank.

I would also like to thank Michael Anderson, Stefano DellaVigna, Frederico Finan, Shachar Kariv, Ethan Ligon, Jeremy Magruder, Craig McIntosh, Matthew Rabin, Emmanuel Saez, Adam Szeidl, and Brian Wright, for their valuable comments and suggestions. I have also benefited from comments received during various seminars at UC Berkeley and other institutions.

I thank the People's Insurance Company of China for their close collaboration at all stages of the project, and I would especially like to acknowledge the contributions of Aijun Cai, Leilao Chen, Baohua Deng, Xiaoping Fan, Sidong Li, Xiangli Li, Zhihua Qian, and Zhanpeng Tao in experiment implementation and data collection. Financial support from 3IE and ILO is greatly appreciated.

Above all, I would like to thank my parents for their constant support and encouragement. They mean the world to me and this undertaking would not have been possible without their love. I must also thank my loving fiancee, Haonan Qu, whose guidance and encouragements have helped me so much get through difficult times.

Chapter 1

Social Networks and the Decision to Insure

1.1 Introduction

It is well recognized that social networks play a fundamental role in shaping human behavior and have important effects on economic outcomes[1]. But how large are social network effects in practice, and what are the underlying mechanisms? In this paper, I investigate this topic in the context of the introduction of a new weather insurance product in rural China. This kind of insurance product is important for households in agricultural communities where production is exposed to substantial weather shocks[2]. However, evidence from several countries shows that participation rates are sub-optimally low, even with heavy government subsidies[3]. One approach to increasing take-up might be to exploit social network effects, which can be an important factor in decisions on whether or not to adopt new and unusual financial products: people may learn about benefits of a product from their friends, imitate their decisions, or respond to their experience.

With data from a two-year randomized experiment in rural China, which includes 185 villages and about 5300 households, I study the influence of social networks on the demand

[1]Existing studies have linked social networks with activities and outcomes including risk sharing (Ambrus et al. (2010)), voting (Lazarsfeld (1944)), finding jobs (Munshi (2004)), trust (Karlan et al. (2009)), financial decision-making (Duflo and Saez (2003); Hong et al. (2004)), technology adoption (Conley and Udry (2010)), etc. For comprehensive discussion and references, see Jackson (2010).

[2]Formal insurance markets are important because informal insurance mechanisms cannot effectively reduce the negative impacts of regional weather shocks, and they leave consumption susceptible to covariate shocks (Townsend (1994)). The absence of insurance markets can lead to highly variable household income and persistent poverty (Dercon and Christiaensen (2011); Jensen (2000); Rosenzweig and Wolpin (1993)).

[3]For example, Giné et al. (2008) found a low take-up (4.6%) of a rainfall insurance policy among farmers in rural India in 2004, while Cole et al. (2011) found an adoption rate of 5%-10% of a similar insurance policy in two regions of India in 2006.

for insurance, both in the short term and over longer periods[4]. The product I am studying is a new weather insurance policy for rice farmers offered by the People's Insurance Company of China (PICC). The novel experimental design allows me not only to identify the causal effect of social networks, but also to test various mechanisms through which social networks operate, including learning about the function and benefits of insurance products, scale effects, imitation, informal risk-sharing, and learning from friends' payout experiences. Additionally, taking advantage of the substantial variation in network structure across villages and households, I am able to test the effect of social structure and initial conditions on the strength of social network effects. Furthermore, using household level price randomization, I calculate the price equivalence of the social network effect.

I begin by investigating the value of social networks for insurance take-up using the experiment implemented in year one. Specifically, if only a subset of farmers were offered financial education about the product, would this have a spillover effect on untreated farmers? To establish causality, I randomly assigned households in each village to early versus late round sessions, and simple versus intensive (including financial education) sessions. For each household, the social network variable is defined as the fraction of a group of friends (named in a social network survey) who were invited to an early round intensive session. I find that, while financial education raises take-up by 43% in the first round, for second-round participants, having one additional friend who obtained first round financial education increases take-up by half as much. I use a household level price randomization to show that this effect is equivalent to decreasing the average insurance premium by 12%.

The magnitude of the social network effect may depend on social structure (Jackson (2008); Jackson and Yariv (2010)) as well as on the initial conditions (Banerjee et al. (2011)). By exploiting variations in village and household level network characteristics, I show that the effect is larger in more clustered villages where households are more interwoven with each other, and when the first set of people to be financially educated are more important in a network sense (they have higher eigenvector centrality). Moreover, households which are more frequently named as friends by other people (higher in-degree), which can more easily be reached by others (smaller length of in-path), or which are more central in the village network, are less likely to be influenced by other people.

After observing a large and significant effect of social networks, it is natural to ask what information conveyed by social networks drives this large effect. Do social networks matter in insurance adoption because they can diffuse knowledge among farmers about the product benefits? Or is it because farmers learn about each other's purchase decisions through social networks and make their own decisions based on that? I find that there is something special about social networks in rural communities: they do not convey information about what other people do, even though others would like to obtain such information, but they do

[4]Most papers are focused on studying ways to improve the initial participation rate, but there is no rigorous study about how to maintain high take-up rates over time as subsidies are gradually removed, which is essential for generating a substantial welfare impact on households.

effectively convey information about what other people know. This result is reached in the following manner. First, I compare the effect of financial education on insurance take-up and knowledge between the two rounds. I find that in the second round, the effect of financial education is smaller, and that farmers understand insurance benefits better when they have a greater number of friends who received financial education. This means that there was diffusion of insurance knowledge from first-round educated farmers to second-round participants. Second, I exploit the exogenous variation in both the overall and individual take-up decisions generated by randomized default options to estimate whether or not subjects are affected by their peers' decisions. No significant effect is found, but, surprisingly, when I told farmers about other villagers' decisions, it actually mattered a lot to them. This suggests that in this case the main mechanism of the social network effect is social learning about insurance benefits, as opposed to a scale effect (greater expected leverage over the insurance company), imitation, or informal risk sharing.

While improving the initial participation rate is crucial as it is a prerequisite for scaling up the program, identifying financially sustainable ways of maintaining high take-up rates over time is also essential. Social networks can be effective for this, because individuals use not only their own experience with the product, but also the experience of others in making purchase decisions. For this, I followed up one year later with a subsample of households from the first year experiment. The randomization in year two includes a household level insurance price. I find that households are not influenced by their friends' purchase behavior during the previous year. However, observing an above median share of friends receiving payouts improves second year take-up at all price levels (by 21.7 percentage points on average) and makes people less sensitive to price change (offsets the price effect by more than 50%). The effect is equivalent to reducing the average insurance premium by 35%. Comparing the effect of learning from others with learning from one's own experience, I show that the observation of friends receiving payouts has about 54% of the impact of a subject's receipt of an actual payout. This means that social networks affect insurance take-up over time through social learning about friends' experience with payouts.

The findings in this paper speak to the commonly used policy of providing a one-time subsidy to a subset of potential buyers with the expectation that this will affect the take-up of the others. My result suggests that such a policy alone cannot lead to indefinitely sustainable voluntary purchases over time. However, combining it with social norms marketing, which disseminates information about the actions of a population's peers, may improve take-up significantly. Moreover, providing financial education to a subset of households and relying on social networks to extend the effect, as well as disseminating payout information effectively, would be good ways of improving the sustained insurance demand.

This paper contributes to the existing literature in the following ways. First, it provides insights on the social network literature. The estimation of the causal effect of social networks is made challenging by the problem of correlated unobservables (Manski (1993); Manski

(1995))[5]. Economists have used experimental approaches and other identification strategies to solve this issue, and the results vary greatly with the context and the product considered[6]. This paper uses randomized experimental methods to estimate the causal effect of social networks, and the monetary equivalence of this effect, in a previously unexplored field. Moreover, while the study of mechanisms through which social networks affect behavior is crucial from both theoretical and policy perspectives, only a few studies to date have shed light on this point[7]. This paper contributes by digging deeper into this question and using experimental designs to directly identify a comprehensive set of general channels of social network effects.

Second, the paper sheds light on the puzzle of how to improve weather insurance take-up. Although existing works have tested the possible explanations of trust, credit constraints, and ambiguity aversion (Giné et al. (2008); Cole et al. (2011); Bryan (2010)), insurance demand is still low even after some of the above treatments were implemented. I provide evidence that a previously unexplored determinant, social networks, can improve both initial participation rates and longer term insurance demand, and significantly reduce households' sensitivity to the removal of subsidies.

Third, this paper contributes to the literature on financial education. 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 (2009a); Lusardi and Mitchell (2007); Stango and Zinman (2009)), the experimental evidence on financial education is mixed[8]. In a context where insurance is new, and farmers have relatively low levels of general education, my results show that lack of financial education is a major constraint on the demand for insurance, and that modest financial training can improve take-up rates significantly.

The rest of the paper is organized as follows. Section 1.2 describes the background for the study and the insurance contract. Section 1.3 explains the experimental design. Section

[5]For example, social norms, homophily, etc.

[6]Papers using a randomized experimental approach to identify the causal effect of social networks include Duflo and Saez (2003), Duflo et al. (2008), Dupas (2010), Kremer and Miguel (2007), Kremer and Levy (2008), and Oster and Thornton (Forthcoming). Examples that use non-experimental methods to identify the effect include: Bandiera and Rasul (2006), Banerjee et al. (2011), Conley and Udry (2010), Foster and Rosenzweig (1995), Munshi (2003) and Munshi (2004).

[7]In general, the mechanisms through which social networks affect the adoption of new technology and financial products include social learning about product benefits, imitation, learning how to use the product, etc. The following two papers are examples which have worked on the mechanisms of social network effects. Kremer and Miguel (2007) find negative peer effects, which effectively rules out explanations such as imitation and learning how to use the product. Banerjee et al. (2011) take advantage of differences in predicted patterns of behavior as a function of the diffusion of basic information (awareness of microfinance) as compared to peer effects (influence of others' decisions). They find that the diffusion of information in microfinance participation is more in line with peer effects than with basic information diffusion.

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

1.4 presents the results, and section 1.5 concludes.

1.2 Background

Rice is the most important food crop in China. Nearly 50% of the farmers produce rice, and more than 60% of the Chinese people use rice as their staple food. In order to maintain food security and shield farmers from negative weather shocks, in 2009 the Chinese government designated the People's Insurance Company of China (PICC) to design and offer the first rice production insurance program to rural households in 31 pilot counties[9]. The program was extended to 62 counties in 2010 and 99 counties in 2011. The experimental sites for this study are two rice production counties included in the second round pilots in Jiangxi province, which is one of China's major rice bowls[10]. In these two counties, rice production is the main source of income for most farmers. Since the product was new, no households had heard of or purchased such insurance before, and most of them had never interacted with PICC before. As a result, farmers, and even government officials at the village or town level had very limited knowledge of weather insurance products and the insurance company as the provider.

The insurance contract is as follows. The actuarially fair price is 12 RMB per mu per season[11]. The government gives a 70% subsidy on the premium, so farmers only pay the remaining 3.6 RMB per mu. If a farmer decided to buy the insurance, the premium will be deducted from the agricultural card, no cash payment is needed[12]. The insurance covers natural disasters including heavy rain, flood, windstorm, extremely high or low temperature, and drought. If any of the above natural disasters happened and led to a 30% or more loss in yield, farmers were eligible to receive payouts from the insurance company. The indemnity rule is illustrated in Figure A.2. The amount of payout increases linearly with the loss rate in yield, with a maximum payout of 200 RMB. The loss rate in yield is investigated and determined by a group of insurance agents and agricultural experts[13]. The average gross income from cultivating rice in the experimental sites is between 700 RMB and 800 RMB per mu, and the production cost is around 300 RMB to 400 RMB per mu. Thus, this

[9]Although there was no insurance before 2009, if major natural disasters happened, governments issued subsidies to households whose production was seriously hurt. However, the level of subsidy was usually very limited and far from enough for farmers to restart production.

[10]The location and map of Jiangxi province is presented in Figure A.1.

[11]1 RMB = 0.15 USD; 1 mu = 0.067 hectare. In the experimental sites, farmers produce two or three seasons of rice each year.

[12]Starting in 2004, the Chinese government has provided rice production subsides to rice farmers in order to give them more production incentives. Each year, subsidies are deposited directly to the agricultural card in the rural credit cooperatives (the main rural bank of China).

[13]For example, consider a farmer who has 5 mu in rice production. If the normal yield per mu is 500kg and because of a windstorm, the farmer's yield decreased to 250kg per mu, then the loss rate is 50% and he will receive 200*50% = 100 RMB per mu from the insurance company.

insurance program provides partial insurance that covers 25% - 30% of the gross income or 50% - 70% of the production cost.

It is also important to note that the post-subsidy price is below the actuarially fair price according to the following calculation. The profit of the insurance company equals revenue minus payouts and fixed cost:

$$\pi = N * Premium - N * P * Payout - FC$$

Where P is the probability of future disasters, and N is the number of households that buy the insurance. Based on communications with local government officials and farmers, the actual probability of disasters that can cause 30% or more loss in yield is estimated to be around 12%. Since N*3.6 $< N$*12%*(200*30%), the post-subsidy price is below fair price. This implies that, as the expected benefit of purchasing insurance is positive for farmers, it is optimal for all farmers who cultivate rice to purchase it. However, as the pre-subsidy price is higher than the fair price, the insurance company earns a profit if the fixed cost is not too large.

1.3 Experimental Design and Data

I use a two-year randomized experiment to identify the role of social networks in influencing insurance demand. The first year experiment was carried out in spring 2010, and includes 185 villages with around 5332 households[14]. The first year data is used to test how social networks affect insurance take-up in the short-run. The second year experiment was implemented in spring 2011; in that phase, I followed 1871 households that were covered in year one, in order to examine social network effects overtime.

1.3.1 Year One: Identify the Short-run Social Network Effect

The first year experiment is used to answer three questions: first, what is the causal effect of social networks on the initial year insurance demand? Second, what is the monetary equivalence of the social network effect? Third, what kind of information did social networks convey that is driving the effect?

In order to generate household level variation in the knowledge of insurance products, two types of information sessions were offered to different households in this year. The simple session took around 20 minutes, during which we only introduced the insurance contract[15]. The intensive session took around 45 minutes and covered all information provided during

[14]In this experiment, "village" refers to the "natural village" in rural China, which is a smaller unit than "administrative villages".

[15]The simple session introduces terms in the contract including the insurance premium, the amount of subsidy provided by the government, the responsibility of the insurance company, the maximum payout, the period of responsibility, rules of loss checking, and the procedures for making payouts.

simple sessions, plus financial education to help farmers understand how insurance works and the product benefits[16].

In each village, two rounds of sessions were offered to introduce the insurance program. During each round, there were two sessions, one simple and one intensive. The second round sessions were held three days after the first round was finished, in order to allow late participants to communicate with early participants. Each household was invited to only one of these four sessions. The effect of social networks on insurance take-up is thus identified by looking at whether second round participants are more likely to buy insurance if they have more friends exposed to financial education in first round intensive sessions.

The experimental design in year one is illustrated in Figure 1.1 and 1.2. There are four randomizations in this experiment, two on the household level and two on the village level. The within-village randomization is presented in Figure 1.1. First, households were randomly assigned to one of the four sessions: first round simple (T1), first round intensive (T2), second round simple (T3), and second round intensive (T4)[17]. This randomization is used to account for exogenous variations among second round participants in the proportion of friends exposed to first round financial education, and hence helps identify the causal effect of social networks within villages.

Second, for each second round session, after the presentation and before participants made final decisions, I randomly divided them into three groups and led different groups to separate rooms, and then disseminated different additional information to the different groups of participants. Specifically, I did not give farmers in groups U1 and U4 any additional information but directly asked them to make take-up decisions; thus, farmers in these two groups received exactly the same information from us as those in the two first round sessions (T1 and T2). For farmers in groups U2 and U5, I told them the overall take-up rate at the two first round sessions in the village that they belong to, so that they knew how many people had attended previous sessions and how many of them had purchased the insurance. For groups U3 and U6, I showed farmers the detailed decision list made in the two first round sessions, so that they knew specifically who purchased the insurance and who did not. This randomization will help determine the main mechanism that drives the social network effect.

Above the within-village randomization, there are two village-level randomizations, which

[16]Sample topics included in the financial education included: How does the insurance program differ from a government subsidy? How much payout can you get under different conditions? What is the expected benefit of purchasing insurance for five continuous years depending on different disaster frequencies and levels, in order to see whether you can gain or lose from buying insurance?

[17]Before doing randomizations, I first approached leaders of each village to obtain a household list that includes name of the household head and basic household characteristics. Households who did not grow rice were excluded. For all household level randomizations in this experiment, I stratified the sample within each village according to household size and area of rice production per capita, and random assignments of households to different treatment groups were made in each stratum. Only household heads were invited to attend one of the four sessions. No one could attend more than one session. In order to guarantee a high session attendance rate, I gave monetary incentives to village leaders and asked them to inform and invite household heads to attend these sessions.

is shown in Figure 1.2. First, I randomly divided villages into two types. In type I villages, all households face the same final price 3.6 RMB, while in each type II village, I randomly assigned seven different final prices ranging from 1.8 RMB to 7.2 RMB to different participants in second round sessions[18]. Type II villages are used to measure the monetary value of the social network effect. The second village-level randomization is only within type I villages, where there was no price variation; in type I villages, I randomized the default option in first round sessions. If the default was BUY, then the farmer needed to sign off if he or she did not want to purchase the insurance; if the default was NOT BUY, then the farmer had to sign on if he or she decided to buy the insurance[19] . Default options were the same in the two first round sessions within each village. The objective of offering different default options was to generate exogenous variations in the first round insurance take-up across villages, which will be used in some estimations as an instrumental variable for first round purchase decisions[20] .

In all cases, households make take-up decisions individually at the end of our visit. Moreover, all households were asked to respond to a household survey, which will be illustrated in detail later.

1.3.2 Year Two: Identify the Social Network Effect Over Time

One year after the first time provision of the rice insurance program, a follow-up experiment was conducted in a subset of villages (72 out of 185, around 2000 households) covered in year one. Both buyers and non-buyers in these villages who were surveyed in year one were sampled for the follow-up experiment. Combining data for both years, I test the effect of social networks on the second year insurance demand curve.

In year two, I randomized the household level subsidy. The randomization strategy is similar as the design in Dupas (2010). In total, eight different prices were offered, with the level of subsidy ranging from 90% to 40%. However, within each village, only two or three prices were assigned[21]. To randomize price sets on the village level, the sample of villages

[18]In all type II villages, farmers in second round sessions T3 and T4 receive exactly the same information as households in first round sessions T1 and T2, respectively. No additional first round take-up information was provided after the presentation.

[19]During sessions where default = BUY, before insurance agents asked farmers to make decisions, instructors told them the following: "We think that this is a very good insurance product, and we believe that most farmers will choose to buy it, so it is more convenient for us to record who does not buy it rather than who buys it. So if you have decided to buy the insurance, there is nothing you need to do, as the premium will be deducted automatically from your agricultural card; if you do not want to buy it, then please come here and sign."

[20]According to Beshears et al. (2009), default options can influence households' financial decisions significantly. A possible reason is that households find it too complicated to make a decision by themselves, so they simply follow the default option as they think it is set as the default because it is a good choice. For more details, refer to Beshears et al. (2009).

[21]Price sets with either two or three different prices were randomly assigned on the village level. For

was stratified according to village size (total number of households) and first year village-level payout ratio. For price randomization on the household level, the sample within each village was stratified according to rice production area. Except for the final price, everything else remained the same in the contract as in year one. With the price randomization, I can test the medium-run social network effect on both the level and the slope of the insurance demand curve.

The procedure is as follows. In each village, I gather households assigned with the same final price and hold meetings for different price groups simultaneously. Households make second year purchase decisions individually right after the meeting[22]. During the meeting, insurance agents briefly repeat items in the insurance contract and announce the list of people in the village who purchased insurance and have received a payout last year, so all households know who in the village received a payout and the amount of the payout[23]. This helps me test the effect of an important type of social network effect: social learning about friends' experiences. Specifically, I calculate the fraction of a group friends (listed in the year one social network survey) who purchased insurance and received a payout; I use that fraction as the measure of social learning about friends' experiences, and test its impact on the second year demand curve.

1.3.3 Data, Summary Statistics and Randomization Check

At the end of the visit in both years, a census was collected in all villages included in the experimental sites. In total, 5332 households were surveyed in year one and 1871 households were surveyed in year two. The survey consists of two components: a household survey and a social network survey.

The household survey contains five parts. The first part asks about household characteristics including household size; age and education of the household head; area of rice production; yields and sales; household income from different sources; borrowing; etc.; The second part asks about types of natural disasters experienced, loss rate in rice yield in the past three years, and methods of coping with such losses. The third part covers experience in purchasing any kind of insurance, as well as payouts received in the past three years. The fourth part asks about risk attitudes and perceptions about future disasters[24]. The

villages assigned with two prices {P1, P2}, P1 <= 3.6 and P2 > 3.6; for villages with three prices {P1, P2, P3}, P1 < 3.6, P2 = {3.6, 4.5}, and P3 > 4.5.

[22]Note that the price experiment is different between the first and second year experiment. In year one, I did price randomization in both first and second round sessions, so second round participants knew different households are facing different prices when they made take-up decisions. However, in year two, because meetings with households in different price groups were held simultaneously, and farmers made decisions at the end of the meeting, they did not know prices are different from household to household before making second year purchase decisions.

[23]After the insurance was offered in April 2010, a low temperature disaster happened in October 2010, just before the harvest of the late season rice, which lead to yield loss for most farmers.

[24]Risk attitudes were elicited by asking sample households to choose between increasing amounts of certain

fifth part contains questions which test farmers' knowledge of how insurance works and its potential benefits; households' trust of the insurance company regarding loss checking and the payout issuing process; and willingness to pay for the insurance. Summary statistics of selected variables are presented in Panel A of Table 1.1. Household heads are almost exclusively male, and the average education level is between primary and secondary school. Rice production is the main source of household income, accounting for 73% of total income on average. Most households have experienced some types of natural disasters in the most recent year, and the average loss rate is around 28%. Households are risk averse on average, and they are more risk averse in the second year compared with in the first year, which can be due to experience of weather shocks during the first year.

The social network component of the survey asked the household head to list five friends, either within or outside the village, with whom the respondent most frequently discusses rice production and financial related problems[25]. The respondent was asked to rank these friends based on which one would be consulted first, second, etc. Relationships with each person named, topics that they usually talk about, and contact frequency were also elicited. I use this data to construct an undirected measure of social links: two households have a link if at least one of them named the other. This measure is used in most of the following estimations and is henceforth called the general measure. In addition, another two measures of social links are defined. One is the strong measure where two households have a link if each of them named the other, and the weak measure where two farmers have a link if at least one of them were named by a friend of the other (second order links). A sample within-village social network map is presented in Figure A.3[26]. In all cases, the social network variable is defined as the fraction of the group of friends with whom they have social links who were invited to a first round intensive session (where financial education was provided). According to Panel B in Table 1.1, the average number of friends listed by households is 4.9, which means only a small proportion of households listed less than five friends. On average, 16% of a household's friends named in the social network survey were invited to the first round financial education provided in the intensive session.

Since how well and how fast information can be diffused among households may depend on network structure such as village size, positions of households in the network, etc., I constructed several measures of village and household level social network characteristics. Summary statistics of these variables are listed in Panel C in Table 1.1. The villages have an average of 32 households. Clustering rates are 0.18. This means that about 20% of

money (riskless option A) and risky gambles (risky option B) in Table A.1. The number of riskless options was then used as a measure of risk aversion. The perceived probability of future disasters was elicited by asking "what do you think is the probability of a disaster that leads to more than 30% loss in yield next year?"

[25]Respondents can list any people except for their parents and children, because in many cases, parents and children cultivate the same plots of rice together.

[26]In Figure A.3, for household 11122066: household 11122076 is an undirected social link, household 11122098 is a strong social link, and household 11122103 is a weak social link.

the time that some household i has connections to some households j and k, j and k also have connections to each other. On average, a household has been named by about three other households. The average path length is about 3.5, which means a household can be connected to any others in the village by passing three households. Eigenvector centrality is a recursively defined notion of centrality: a household's centrality is defined to be proportional to the sum of its friends' centrality. The larger the centrality, the more important is the person's position in the village network.

The overall take-up rate in the first year is 44%, while that in the second year is higher, about 50%. According to Panel D and E in Table 1.1, around 56% households who purchased the rice insurance received some payouts during the first year, and the average amount of payout was about 86 RMB per mu.

Randomization checks are presented in Table A.2 and A.3. According to Table A.2, most household characteristics are balanced across the four different sessions in the first year experiment. To check whether the price randomization is valid, I estimate the following regression in both of the type II villages where price variation was implemented in year one, and in all villages in year two:

$$X_{ij} = \alpha_0 + \alpha_1 Price_{ij} + \alpha_2 Price_{ij}^2 + \eta_j + \epsilon_{ij} \tag{1.1}$$

Where X_{ij} is a set of characteristic of household i in village j, including gender, age, household size, education, and area of rice production. $Price_{ij}$ is the post-subsidy price faced by household i in village j, and η_j includes village dummies. Estimation results are listed in Table A.3, showing that none of the coefficient are statistically significant. As a result, the price randomization was valid in both years.

1.4 Estimation Strategies and Results

1.4.1 Effect of Social Networks on Initial Year Insurance Demand

This section tests the effect of social networks on insurance demand when it was offered at the first time, and explains what types of information act as the mechanism driving the effect, using data collected in year one.

Estimation of the Social Network Effect

The average take-up rate in different information sessions is illustrated in Figure 1.3. It shows that, while the difference between the two first round sessions is substantial, there is almost no difference between the two second round sessions. Moreover, the take-up rate of second round sessions is much higher than that of first round simple sessions. The above evidence suggests that the financial education provided during first round intensive sessions improved farmers' take-up rates, and that, during the three days' time interval between the

two rounds of sessions, there was information diffusion from first round to second round participants.

To estimate the effect of social networks on insurance take-up, using the experimental design in Figure 1.2, I use only type I villages (those without price randomization). First, I estimate the financial education effect using the sample of first round participants (simple session T1 vs. intensive session T2) using the following equation:

$$Takeup_{ij} = \beta_0 + \beta_1 Intensive_{ij} + \beta_2 X_{ij} + \eta_j + \epsilon_{ij} \tag{1.2}$$

where $Takeup_{ij}$ is an indicator of the purchase decision made by household i in village j, which takes a value of one if the household decided to buy the insurance and zero otherwise. $Intensive_{ij}$ is a dummy variable equal to one if household i was invited to one of the two intensive sessions in village j and zero otherwise. X_{ij} includes household characteristics such as gender, age, production size, etc., and η_j includes village dummies. According to results in Table 1.2, the take-up rate of first round intensive sessions (50%) is 15 percentage points higher than that of first round simple sessions (35%), suggesting a large and significantly positive financial education effect that increases take-up rate by 43% in the first round.

Second, to test the social network effect, i.e. the spillover effect of first round financial education on second round participants, I focus on the sample of households who were assigned to second round groups U1 and U4 (where no first round take-up information was revealed) and test whether they are more likely to buy insurance if they have more friends who were invited to the first round intensive session (which included financial education):

$$Takeup_{ij} = \tau_0 + \tau_1 Network_{ij} + \tau_2 X_{ij} + \eta_j + \epsilon_{ij} \tag{1.3}$$

where the social network measure is defined as the fraction of the group of friends named in the social network survey who have been invited to a first round intensive session (the general social network measure)[27]. Since households are more likely to be exposed to the information provided during financial education (which improves insurance take-up as shown in Table 1.2) if more of their friends attended a financial education session, a positive social network effect is expected.

Estimation results are listed in Table 1.3. According to column (1), I find a significantly positive effect of social networks on insurance take-up, and the magnitude of that effect is around 33.7 percentage points. This suggests that having one additional friend attending a first round intensive session - in other words, raising the general network measure by 20% because each household lists five friends - increases a farmer's own take-up rate by 33.7 * 0.2 = 6.74 percentage points. This explains about 45% of the direct financial education effect (column (1) in Table 1.2). In column (2), some control variables were added into the

[27]For example, if a household listed five friends, and two of them were invited to a first round intensive session, then the social network measure equals 0.4.

regression[28]. This column shows that the magnitude of the social network effect does not change much with additional controls. In addition, it suggests that older people, farmers with larger production size, or those with more education are more likely to buy the insurance. Moreover, households who are more risk averse, or those who predict a higher probability of natural disasters in the following year, are more likely to purchase insurance. In column (3), I test whether the magnitude of social network effects depends on whether a farmer received financial education. The social network effect is larger in second round simple sessions, according to the significantly negative coefficient of the interaction term between the network measure and intensive session. This means people are less influenced by their friends when they have a better understanding of the product.

The magnitude of social network effects may depend on the strength of ties[29]. To test this, I extend the framework to include different measures of social links: the strong measure (mutually named links) and the weak measure (second order links), and re-estimate equation (1.3) using these two measures. Results are listed in Table 1.4. According to column (2), having one additional strongly linked friend attending first round financial education improves a farmer's own chance of taking the insurance by 8.5 percentage points, which is larger than the effect of undirected social links (6.7 percentage points). In contrast, friends with weak links are much less influential: according to column (4) in Table 1.4, the number of weakly linked friends receiving first round financial education does not have a significant effect on a farmer's own behavior. This means that during a short period of time (three days in this case), households are not influenced much by friends' friends. In addition, I test whether the magnitude of social network effects varies according to the relationship between the farmer and the person he or she named in Table 1.5. The results show that government officials have the largest effect on influencing their friends' purchase decisions, followed by relatives and neighbors.

While having friends exposed to financial education improves a farmer's own take-up, the effect can be nonlinear. The question is whether having just one friend receiving financial education is enough, or having more friends receiving financial education generates larger effects. I test the nonlinearity of the social network effect using the following equation:

$$Takeup_{ij} = \rho_0 + \rho_1 One_{ij} + \rho_2 Two_{ij} + \rho_3 More_{ij} + \rho_4 X_{ij} + \eta_j + \epsilon_{ij} \tag{1.4}$$

where One_{ij}, Two_{ij}, and $More_{ij}$ are dummy variables which equal one if household i has one, two, or more than two friends assigned to the first round intensive session (the session with financial education) in village j, and 0 otherwise. Results are presented in Table 1.6.

[28]Because a small proportion of households named less than five friends in the social network survey, and these households might be different from other farmers in some aspects, I did a robustness check by excluding these households and re-estimating the social network effect. Results in Table A.4 show that the significance and magnitude of the social network effect remained almost the same.

[29]There are various ways to measure the strength of a tie, and in some contexts, effect of weak ties on outcomes can be as important as that of strong ties Granovetter (1973).

According to column (1), the magnitude of the social network effect increases with the number of friends receiving first round financial education. Specifically, among second round participants, the effect on insurance take-up of having two friends obtaining first round financial education is 20.2 percentage points; this is about 14 percentage points higher than the effect of having only one friend financially educated in the first round, which is only 5.6 percentage points. However, having more than two friends financially educated has just a slightly higher effect on take-up (8 percentage points higher) than that of having two. The effect holds when additional controls are added (column (2)). As a result, the social network effect increases with the number of friends receiving financial education, but the marginal return is diminishing when the number is larger than two.

Social Network Effect, Network Characteristics, and Injection Points

How efficiently information can be diffused, and to what extent a farmer can be influenced by other farmers, may depend on network characteristics at both the village and individual level. First, the role of village-level network structure is obviously important in most diffusion models[30]: for example, the spread of information might be easier and faster in more clustered villages where households are more strongly connected with each other. Taking advantage of cross-village variations in network structure, I estimate how the magnitude of social network effect varies with a set of variables that capture village-level network structure, including: village size, clustering, segmentation, etc., using the following regression:

$$Takeup_{ij} = \eta_0 + \eta_1 Network_{ij} + \eta_2 VilNetCharact_j + \eta_3 Network_{ij} * VilNetCharact_j + \eta_4 X_{ij} + \epsilon_{ij} \tag{1.5}$$

where $VilNetCharact_j$ includes different measures of village-level network characteristics (detailed definitions in Appendix A.2). According to the results in table 1.7, while characteristics such as village size, transitivity, etc. do not affect the magnitude of the social network effect, clustering does: according to column (2), in villages with a higher clustering coefficient (households are more strongly connected with each other), the effect of social networks on insurance take-up is larger, suggesting that, in these villages, information diffusion is more efficient.

Second, the "injection points" (the households that were included in the first round financial education) could affect diffusion in a social network[31]. The question I want to answer is: in a village with hundreds of people, if only a small group of households can be formally financially educated, will the strength of the social network effect depend on which individuals are initially contacted, as well as on the importance in the village network of farmers who make the take-up decision? I define a set of individual level network characteristics, including

[30] See Chapter 7 in Jackson (2008) for detailed discussion and background.

[31] The idea that injection points may matter has been discussed in Katz and Lazarsfeld (1955), Ballester et al. (2006), and Banerjee et al. (2011). However, most of the previous literature has not tested that explicitly, with the exception of Banerjee et al. (2011).

in-degree (how often is a farmer named by other households), path length (the average length of the path between a farmer and any other households), and eigenvector centrality, and answer this question by estimating the following equation[32]:

$$\begin{aligned} Takeup_{ij} = \eta_0 + \eta_1 Network_{ij} + \eta_2 NetCharact_{ij} + \eta_3 Network_{ij} * NetCharact_{ij} \\ + \eta_4 OwnCharact_{ij} + \eta_5 Network_{ij} * OwnCharactij + +\eta_6 X_{ij} + \eta_j + \epsilon_{ij} \end{aligned} \tag{1.6}$$

where $NetCharact_{ij}$ is the average network characteristics of friends named by household i who attended the first round intensive session in village j, and $OwnCharact_{ij}$ is the network characteristics of household i. The coefficient η_3 tells us what types of person have more influence on others, while coefficient η_5 suggests who is more likely to be influenced by others. According to the results in Table 1.8, the average in-degree or out-path length of one's friends does not influence the magnitude of the social network effect, but their eigenvector centrality is. Referring to column (6), if the eigenvector centrality of the set of friends in first round financial education is one standard deviation larger (0.1), second round overall take-up is around 5 percentage points larger, and the effect of social networks (having one additional friend financially educated) on take-up is around 6.4 percentage points larger. Another noteworthy point is that what matters more is a farmer's own network characteristics: according to columns (2), (4) and (6), those who were named more often by other households (higher in-degree), who can be reached more easily (smaller in-path length), and who have a more important network position (higher eigenvector centrality), are less likely to be influenced by other people. However, these characteristics only affect the magnitude of the social network effect, and do not directly affect the take-up decisions.

Given that individual-level network characteristics are important, what are the determinants? I show this in Table A.5. According to column (4) in Table A.5, for example, households with higher centrality in the network are usually better educated and have larger production size. The reason that they are more influential may be that they better understand the benefits of purchasing insurance and can explain it better. A reason that they are less likely to be influenced by others may be that, as they are larger farmers, having insurance is more important for them, so they purchase it regardless of what other people say. I present the heterogeneity of social networks depending on these easily observable determinants in Table A.6.

[32]While Banerjee et al. (2011) studied the effect of group-level characteristics of initially treated households on village level microfinance participation, here I define injection point characteristics for each household's social links. I explain how the magnitude of social network effect depends on characteristics of friends who were financially educated first and on each farmer's own importance in the village network.

Monetary Equivalence of the Social Network Effect

In order to understand the importance of the social network effect better, I assess the price equivalence of it using type II villages, where price randomization was implemented. Specifically, I estimate the effect of removing subsidies on insurance demand, and test whether households are less sensitive to subsidy removal if they have more friends exposed to financial education. Then I calculate the price equivalence of the social network effect based on estimated coefficients. I write a simple theoretical model of insurance demand to explain why social networks can potentially influence both the level and slope of insurance demand curve; this is presented in Appendix A.3. The slope of the insurance demand curve can be influenced by household level perceptions of the expected benefit of the new insurance product, and the uncertainty about the expected benefit. The degree of concentration or dispersion in distribution of the expected product benefit can also influence the slope of the demand curve. All of these factors can be influenced by social networks. For example, when a farmer has friends receiving financial education in the first round, the farmer may be less uncertain about the product benefit after friends explained it; as a result, the farmer who learned insurance knowledge from friends may be less sensitive to a price increase than a farmer who did not learn insurance knowledge through any social channels.

In Figure 1.4, I compare the insurance demand curve between households who have an above-median proportion of friends receiving financial education and those who have a below-median proportion of friends who were financially educated. It is clear that the insurance demand curve is higher and flatter, especially under high prices, when a high proportion of friends is exposed to financial education in intensive sessions. To estimate this empirically, I use the following estimation equation:

$$Takeup_{ij} = \gamma_0 + \gamma_1 Price_{ij} + \gamma_2 Network_{ij} + \gamma_3 Price_{ij} * Network_{ij} + \gamma_4 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.7)$$

where $Price_{ij}$ is the final price assigned to household i in village j, which ranges from 1.8 RMB to 7.2 RMB, with a total of seven different prices. The coefficient of the interaction term between $Price_{ij}$ and $Network_{ij}$ tells us whether the social network effect can substitute for part of the subsidy effect. Results are presented in Table 1.9. Column (1) shows that reducing subsidies by 10% (1.12 RMB) decreases take-up by around 10 percentage points. However, according to column (2), the interaction term between social network and price is significantly negative, which suggests that households are less sensitive to subsidy removal if they have more friends receiving financial education. Specifically, having one additional friend receiving financial education mitigates the price effect by 0.13*0.2/0.167 = 16%.

One concern of the above estimation is that, for households in the price experiment, some of their friends face lower prices than they do, while some of their friends face higher prices. How this "fairness" issue might bias the price elasticity of insurance demand is uncertain. To control for this, I included two additional variables when estimating equation (7): %friends with prices higher than one's own price and %friends with prices lower than

one's own price. Referring to the results in column (3) of Table 1.9, the effect changed slightly: having one additional close friend receiving financial education mitigates the price effect by around 0.15*0.2/0.151 = 20%, and is equivalent as the effect of a 12% decrease in the average insurance premium[33].

In summary, the above results tell us that providing financial education about the product improves insurance take-up significantly. More importantly, it has a large and significant spillover effect on insurance adoption by other farmers: among second round participants, having one more friend attending financial education has an effect equal to almost 45% of the first order education effect, and is as effective as reducing the average insurance premium by 12%.

Mechanisms of the Social Network Effect

A natural question is why social networks matter. What did people actually learned from their friends? Understanding these mechanisms is a prerequisite for making policy recommendations as to how governments or institutions can use social networks to obtain efficient outcomes. Generally speaking, social networks may influence the adoption of new technology or financial products because individuals care about other people's decisions (Bandiera and Rasul (2006); Banerjee (1992); Ellison and Fudenberg (1993); Rogers (1995)), or because people learn how to use the product from their friends (Duflo and Saez (2003); Munshi and Myaux (2006); Kremer and Miguel (2007); Oster and Thornton (Forthcoming)). Social networks could also matter if people learn the value or benefits of a product from their friends (Kremer and Miguel (2007); Koher et al. (2001)).

Because insurance is a financial product rather than a technology, people do not need to learn how to use it, and thus this paper does not consider the "learn to use" channel. I will focus instead on the importance of two types of information that can be conveyed by social networks and can influence people's behavior: insurance knowledge and purchase decisions. If the reason that farmers are affected by their friends' exposure to financial education is that their understanding of insurance benefits is improved by learning from their friends, this means that insufficient knowledge of insurance impairs adoption; in that case, providing financial education would be crucial. On the other hand, farmers could be influenced by their friends' behavior in deciding whether to buy insurance. The influence could occur because of scale effects (famers have greater leverage over the insurance company if more of them purchase together), imitation (farmers want to act like each other), or informal risk-sharing (a farmer's decision depends on the purchasing decision of households from which the farmer borrows or to which the farmer lends). If the network effect is driven by friends' purchase

[33]I calculate the price equivalence X of the social network effect by the following formula:

$$X = \frac{\frac{Coef(Network)[0.2*(1-Coef(Price*Network)*Mean(price))]}{Coef(Price)[1-Coef(price*Network)*Mean(Network)]}}{Mean(Price)}$$

decisions, then using low-cost marketing strategies to guarantee a high adoption rate by pilot clients could significantly improve the take-up rate by follow-up customers.

First, to test the "insurance knowledge" mechanism, I use the following two strategies. The first method is to compare the magnitude of the financial education effect on insurance take-up and knowledge between first round (simple session T1 vs. intensive session T2) and second round sessions (simple session U1 vs. intensive session U4, with no take-up information revealed in either U1 or U4). Intuitively, if late participants can obtain insurance knowledge from early participants during the time interval between the two rounds, then they have already been exposed to the information we provided during financial education before attending their own sessions, regardless of whether they were assigned to the simple or intensive session. As a result, we should see a smaller effect of financial education in the second round. Estimation equations are as follows:

$$Takeup_{ij} = \omega_0 + \omega_1 Intensive_{ij} + \omega_2 Second_{ij} + \omega_3 Intensive_{ij} * Second_{ij} + \omega_4 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.8)$$

$$Knowledge_{ij} = \omega_0 + \omega_1 Intensive_{ij} + \omega_2 Second_{ij} + \omega_3 Intensive_{ij} * Second_{ij} + \omega_4 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.9)$$

where $Second_{ij}$ is a dummy variable which equals one if a household is assigned to one of the two second-round sessions in his village, and $knowledge_{ij}$ is a measure of insurance knowledge, which is defined as the score that a household got in the ten questions that we asked during the household survey to test farmers' insurance knowledge. Results are presented in Table 1.10. Column (1) shows that while financial education raises the take-up rate significantly in the first round (14 percentage points), it makes almost no difference in the second round. Similar patterns regarding insurance knowledge can be found in columns (3). Moreover, while the levels of insurance take-up and insurance knowledge between the two second round sessions are very similar, they are significantly higher than that of first round simple sessions.

The second strategy to identify the insurance knowledge mechanism is to test whether households perform better in the insurance knowledge test when they have more friends attending first round financial education, by estimating the following equation:

$$Knowledge_{ij} = \lambda_0 + \lambda_1 Network_{ij} + \lambda_2 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.10)$$

According to column (4) of Table 1.10, having one additional friend assigned to first round intensive sessions improves the level of insurance knowledge by 7 percentage points. Furthermore, I test whether the effect is larger when one's friends better understand the materials provided during financial education, and as a result can better teach other people:

$$\begin{aligned} Knowledge_{ij} = {} & \lambda'_0 + \lambda'_1 Network_{ij} + \lambda'_2 NetKnowledge_{ij} \\ & + \lambda'_3 Network_{ij} * NetKnowledge_{ij} + \lambda'_4 X_{ij} + \eta_j + \epsilon_{ij} \end{aligned} \quad (1.11)$$

where $NetKnowledge_{ij}$ is the average insurance knowledge test score received by household i's friends who obtained first round financial education in village j. Column (5) in Table 1.10 shows that a farmer can learn more insurance knowledge from friends who were financially educated when those friends better understand the information provided in financial education. These results support the argument that, during the three days between the two rounds of sessions, second round participants obtained insurance knowledge from the first set of people who were financially educated, and that such informal learning of insurance knowledge improved insurance take-up significantly.

Second, to estimate whether social networks conveyed the "purchase decision", I directly test the effect of other people's decisions, i.e., overall take-up rate in first round sessions, and friends' take-up rate in first round sessions, on second round participants' behavior. Consider the effect of overall first round take-up first:

$$Takeup_{ij} = \gamma_0 + \gamma_1 TakeupRate_j + \gamma_2 Reveal_{ij} + \gamma_3 TakeupRate_j * Reveal_{ij} + \gamma_4 X_{ij} + \epsilon_{ij} \quad (1.12)$$

where $TakeupRate_j$ is the overall take-up rate in first round sessions (T1 and T2) in village j, which is a continuous variable ranging from 0 to 1, and $Reveal_{ij}$ is an indicator of whether we told second round participants about the overall first round take-up rate. The hypothesis is that individuals are more likely to purchase insurance if they see higher take-up rates in previous sessions, because of either a scale effect or imitation. However, OLS cannot provide a consistent estimates because unobservable variables such as social norms may affect both $TakeupRate_j$ and $Takeup_{ij}$.

As shown in column (1) of Table 1.11, randomized default options in first round sessions give significant and substantial variations in the overall first round take-up rates: the average take-up rate of "default = BUY" sessions is around 12 percentage points higher than that of "default = NOT BUY" sessions. As a result, I can use the default option as the IV for first round overall take-up rates. OLS and IV estimation results for the whole second round sample are presented in columns (2) and (3) of Table 1.11: farmers are more likely to buy insurance when the overall first round take-up rate is higher, but the effect is much smaller if we did not explicitly reveal that information. To see this more clearly, I break down the sample and re-estimate the influence of first round overall take-up rate in columns (4) - (5). According to column (4), second round participants are not influenced by decisions made by first round participants. However, they do care about what other farmers do: as suggested by columns (5), if we disseminated first round overall take-up information during second round sessions, then increasing the first round take-up rate by 10% can raise take-up rate by around 4.3%, which is almost half of the first-order effect.

To see if friends' behaviors have similar effects as overall take-up on farmers' decisions, I estimate the following equation using the sample of second round participants who did not receive take-up information and those who did receive first-round decision list from us (U1,

U3, U4, and U6 in Figure 1.1):

$$\begin{aligned} Takeup_{ij} = \delta_0 + \delta_1 TakeupRate_j + \delta_2 TakeupRateNetwork_{ij} + \delta_3 Reveal_{ij} \\ + \delta_4 TakeupRate_j * Reveal_{ij} + \delta_5 TakeupRateNetwork_{ij} * Reveal_{ij} + \\ + \delta_6 X_{ij} + \epsilon_{ij} \end{aligned} \tag{1.13}$$

where $TakeupRateNetwork_{ij}$ represents the take-up rate among friends of household i who attended first-round sessions in village j[34]. Similar to what has been discussed before, both $TakeupRate_j$ and $TakeupRateNetwork_{ij}$ are endogenous. While I still use first round default option as the IV for the overall first round take-up rate, I use two IVs for $TakeupRateNetwork_{ij}$: Default*%friends in first-round sessions (first round default options are more likely to influence the number of friends who purchase insurance if more friends are included in first round sessions), and %friends who were in the first round and were assigned to the intensive session (because financial education has a positive effect on insurance take-up, friends' take-up rate in the first round should be higher if more of them were assigned to the intensive session[35]). According to results in Table 1.12, decisions made by friends in a farmer's social network do not influence the farmer's own decision (column (4)). This is not because farmers do not care about other villagers' decisions, as this information has a large and significant influence if we explicitly revealed it (column (5)), but because, at least in this context, social networks did not convey such information.

The above results suggest that, while social networks are efficient in transmitting insurance knowledge, they do not convey purchase decisions. This is surprising, because farmers actually care a good deal about that type of information, which we know because of the significant effect if the information is explicitly revealed. I conclude that the short-run social network effect on insurance take-up is mainly driven by the diffusion of insurance knowledge, as opposed to mechanisms related to the behavior of other people, such as scale effect, imitation, or informal risk-sharing.

There are two potential explanations as to why farmers do not tell each other what they do. First, it is possible that it takes time for information about purchase decisions to be revealed and diffused, so it cannot be captured during the three days between the two rounds. I can check this possibility by testing whether farmers are influenced by their friends' decisions over time using second year take-up data. Second, because of cultural factors and the type of financial product they are considering, farmers may be reluctant to reveal their decisions, because they are not sure whether they have made the right decision[36].

[34]For example, if a household named five friends in the social network survey, four of them were assigned to the first round intensive session, and two of them decided to buy the insurance, then $TakeupRateNetwork_{ij} = 0.5$.

[35]If, among the five friends listed, three of them were invited to first-round sessions and two of them were assigned to first-round intensive sessions, then the "fraction of friends in first-round" = 2/5= 0.4, and the "fraction of first-round friends in intensive sessions" = 0.5.

[36]Please note that while first round participants made purchase decisions directly after the meeting, second

Several policy implications can be drawn from the above results. First, providing financial education to a subset of farmers, and depending on social networks to multiply its effect on others, can substantially improve the overall insurance take-up rate. Second, these results cast some doubt on the common practice of providing heavy subsidies for innovative products to a sub-sample of potential customers in order to encourage take-up, with the hope that other people will follow their behavior. The fact that farmers do not communicate purchase decisions tells us that providing subsidies to a sub-population may not be sufficient to achieve the expected outcomes, at least in some contexts. However, social norms marketing, in which we disseminate information to the full population about the behavior of peers, might be effective as a supplemental strategy. Because we do see increased take-up when farmers are explicitly informed about the behavior of others, combining either education or subsidies for a sub-population with social norms marketing may be an inexpensive way to expand the take-up rate for innovative products to a broader population[37].

1.4.2 Effect of Social Networks on Second Year Insurance Demand

My results so far show that social networks significantly improve the first year insurance demand through diffusing insurance knowledge. However, if we cannot maintain voluntary insurance purchase over time when subsidies are gradually removed, the insurance program cannot be financially sustainable or be effective on improving household welfare. As a result, it is important to estimate the social network effect on insurance demand over longer periods. In this section, I estimate the impact of social networks, especially social learning about friends' payout experience, on the second year insurance demand curve, using survey data from both years and payout data from the insurance company[38].

Social Network Effect Over Time

In this part, I estimate the social network effect one year after the insurance was offered the first time. Specifically, I look at the effect of friends' take-up decisions in the previous year on a farmer's own insurance demand curve in the current year. The estimation equation is

round participants did not know that. Therefore, even if they did not reveal the purchase decisions (they can simply say they have not decided yet), they will be not blamed by meant to do that.

[37]Field experiments have shown that social norms marketing which tries to exploit people's tendency to imitate peers, have mixed effect on decision-making (Beshears et al. (2011); Cai et al. (2009); Carrell et al. (2011); Chen et al. (2010); Frey and Meier (2004); Fellner et al. (2011)). However, there is little evidence on how social norms marketing affects choices in fields such as insurance.

[38]According to Figure A.4, the correlation between loss rate in yield and the probability of receiving payouts is consistent with what was stated in the contract.

as follows:

$$Takeup_{ij2} = \sigma_0 + \sigma_1 Price_{ij2} + \sigma_2 NetworkTakeup_{ij1} + \sigma_3 Price_{ij2} * NetworkTakeup_{ij1} + \sigma_4 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.14)$$

where $Takeup_{ij2}$ is a dummy equal to one if household i in village j bought the insurance in year two; $Price_{ij2}$ is the final price faced by household i in village j in year two, which ranges from 1.2 RMB to 7.2 RMB; $NetworkTakeup_{ij1}$ is the proportion of friends who have purchased insurance in year one[39]; X_{ij} includes household characteristics such as age, gender, education of household head, household size, rice production area, insurance knowledge, risk attitude, and perceived probability of future disasters in year two; and η_j includes village dummies.

However, OLS estimation cannot give a consistent estimation of the above equation because friends' decision in year one is endogenous. According to results reported in section 1.4.1, default options and first round financial education substantially influence first-year insurance take-up. As a result, here I use two IVs for $NetworkTakeup_{ij1}$: %friends in first round sessions * default option, and %friends in first round intensive sessions[40].

Table 1.13 presents the estimation results. First, raising prices (removing subsidies) has a significantly negative effect on insurance demand in year two: on average, removing subsidies by 10% decreases insurance demand by around 5.04 percentage points. Second, OLS estimation results in columns (1) and (2) show that farmers are slightly more likely to purchase insurance in year two if more friends purchased it in year one. Specifically, having one additional friend buy insurance in year one increases the probability of purchase in year two by around 2.3 percentage points; this result is barely significant. Turning to the IV result, the first-stage estimation result in column (3) shows that the two IVs affect friends' decisions in year one significantly. However, according to columns (4) and (5), the IV estimation result shows that friends' decisions in year one does not have a significant effect on a farmer's own decision in year two. Moreover, as presented in column (6), the interaction term between price and friends' decisions in year one is also not significant, which means decisions made by friends in year one do not influence either the level or the slope of the insurance demand curve in year two.

In summary, households do not copy their friends' decisions over time. However, this does not mean that farmers are not influenced by their friends at all, because what they observe over time is not only friends' purchase decisions, but also their experience. It is possible

[39]For example, if a household listed five friends during the social network survey in year one, and three of them purchased insurance in year one, then the variable is defined as $3/5 = 0.6$.

[40]One concern of using %friends in first round intensive sessions as the IV is that this may affect both friends' financial knowledge and a farmer's own knowledge through information diffusion in networks, which will further affect a farmer's insurance demand curve in year two. To deal with this problem, I controlled for households' insurance knowledge in year two in this estimation.

that some specific types of experience of their friends still influence their own decisions[41]. The most important type of experience that farmers care about in this context is receiving payouts from the insurance company: after the first year, if disasters happened, farmers can see whether anyone received a payout, whether loss checking and payout issuance are done fairly, etc., which can potentially affect their own long-run insurance demand. As a result, in the next part, I measure the effect of observing friends' payout experience on the second year insurance demand curve.

Learning from Own and Friends' Payout Experience

First, I study the effect of directly receiving payout in the first year on the insurance demand curve in the following year, using the sample of households who purchased insurance in the first year. In Figure 1.5, I compare the insurance demand curve of households that received payout to the insurance demand curve of households that did not. It shows that people are significantly more likely to renew the contract if they received some payouts in the first year. Furthermore, those who received payouts are less sensitive to subsidy removal in the second year; the insurance demand curve is almost flat for them. I then estimate whether these effects are statistically significant using the following regression:

$$Takeup_{ij2} = \phi_0 + \phi_1 Price_{ij2} + \phi_2 Payout_{ij1} + \phi_3 Price_{ij2} * Payout_{ij1} + \phi_4 X_{ij} + \eta_j + \epsilon_{ij} \quad (1.15)$$

where $Payout_{ij1}$ is a dummy equal to one if the household received payout last year. According to results in columns (1)-(3) in Table 1.14, receiving payouts improves second year take-up rate by around 40 percentage points, and mitigates the subsidy removal effect by around 70%. While this effect could be driven by "learning by doing", it could also be explained if people changed their risk altitudes or perceived probability of future disasters after experiencing some weather shocks in year one. In order to control for this effect, I use a regression discontinuity method to re-estimate the effect of receiving payouts on renewal behavior, using loss rate in yield during the first year as the running variable. According to the results in columns (4) and (5) in Table 1.14, the results stay almost the same, so the weather shock mechanism can be ruled out as a possible explanation.

Next, I compare the effect of "learning by doing" to that of social learning about friends' payout experience. By definition, a farmer who did not buy insurance in the first year could not directly receive a payout, so his or her second year behavior cannot be explained as "learning by doing". Regardless of whether a farmer purchased insurance in the first year, observing other people receiving payouts may affect the second year take-up decision, which is a specific social network mechanism which I describe as social learning about friends' payout experience.

In Figure 1.6, I compare the second year insurance demand curve of households who have

[41]Several papers have shown the importance of learning from own or friends' experience in the process of new product adoption (Besley and Case (1994); Conley and Udry (2010); Dupas (2010)).

an above-median proportion of friends receiving payouts in the first year and that of those who have a below-median proportion of friends receiving payouts. We can see that, when a farmer has more friends who received payouts, the insurance demand curve is significantly higher and flatter. To estimate this empirically, I use the following regression:

$$Takeup_{ij2} = \psi_0 + \psi_1 Price_{ij2} + \psi_2 NetworkPayoutHigh_{ij1} + \psi_3 Price_{ij2} * NetworkPayoutHigh_{ij1} + \psi_4 NetworkTakeup_{ij1} + \psi_5 X_{ij} + \eta_j + \epsilon_{ij} \tag{1.16}$$

where $NetworkPayout_{ij1}$ is the proportion of friends in one's social network who have purchased insurance in the first year and received payout[42], and $NetworkPayoutHigh_{ij1}$ is a dummy which is defined as one if $NetworkPayout_{ij1}$ is higher than the sample median and zero otherwise. It is possible that households' insurance demand curve changes according to friends' payout experience because households may update their beliefs about the potential benefits of this product or the uncertainty about this program after observing friends' experience with it.

I estimate equation (1.16) using three different samples: the whole sample, only those who purchased insurance in the first year, and only those who did not purchase in the first year. According to columns (1), (3) and (5) in Table 1.15, overall, households who have an above-median proportion of friends receiving payouts are about 22 percentage points more likely to purchase insurance in year two. However, the effect is only significant for those who did not purchase insurance in the first year; those who bought insurance in year one do not care about whether other people received payout. Moreover, as shown in columns (2), (4) and (6) of Table 1.15, friends' payout experience affects price sensitivity for all households, regardless of whether a household purchased insurance in the first year. Specifically, observing an above-median proportion of friends receiving payouts mitigates almost 50% of the negative subsidy removal effect, which is equivalent to the effect of reducing the average insurance premium by 35%[43].

As a result, the effect of learning from friends' experience on the level of insurance demand equals about half the effect of directly receiving payout; the effect of learning from friends' experience on the slope of the insurance demand curve equals about 70% of the learning by doing effect. Consequently, observing more friends receiving payouts can substantially influence a farmer's own insurance demand and price sensitivity in future periods.

[42]For example, if a household listed five friends in the social network survey, four of them purchased insurance in year one, and two of them received payouts, then the variable is defined as $2/4 = 0.5$.

[43]I calculate the price equivalence of the social network effect X by the following formula:

$$X = \frac{\frac{Coef(NetworkPayoutHigh)[(1-Coef(Price*NetworkPayoutHigh))*Mean(price)]}{Coef(Price)[1-Coef(price*NetworkPayoutHigh)*Mean(NetworkPayoutHigh)]}}{Mean(Price)}$$

.

1.5 Conclusions

This paper uses a two-year randomized field experiment in rural China to analyze social network effects in the adoption of a new weather insurance product. I find strong evidence that social networks play important roles in affecting insurance take-up through social learning mechanisms, in both the short and medium-run. In the first year, providing financial education to a subset of farmers has large and positive spillover effect on other farmers; this is driven by the diffusion of insurance knowledge through social networks, as opposed to diffusion of behavior, which in turn might reflect an effort to gain greater leverage over the insurance company (acquiring strength in numbers), imitation (acting like the others), or informal risk-sharing. One year later, through social learning of friends' payout experiences, social networks affect the insurance demand of households that did not purchase insurance previously. For all households, observing friends receiving payouts significantly reduces the negative effect of subsidy removal. These results suggest that policy interventions such as providing intensive financial education to a subset of households and relying on social networks to multiply its effect on others at the early stage, disseminating information on payouts when they are made, or combining subsidy policies with dissemination of peers' decisions, can remarkably improve the demand for weather insurance.

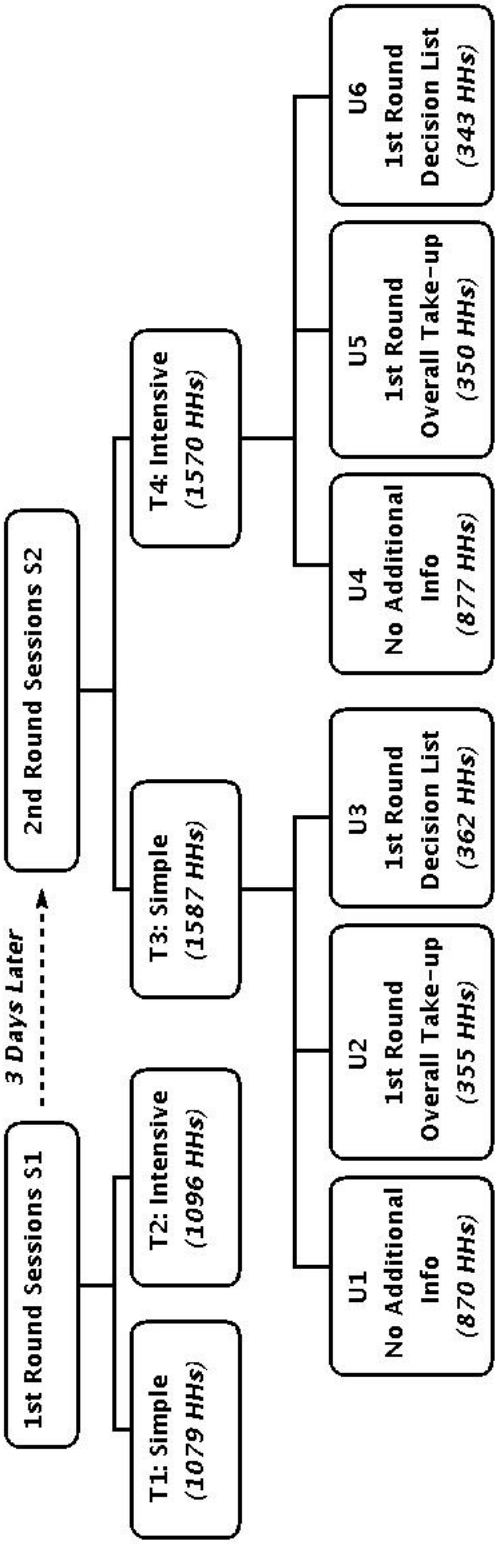
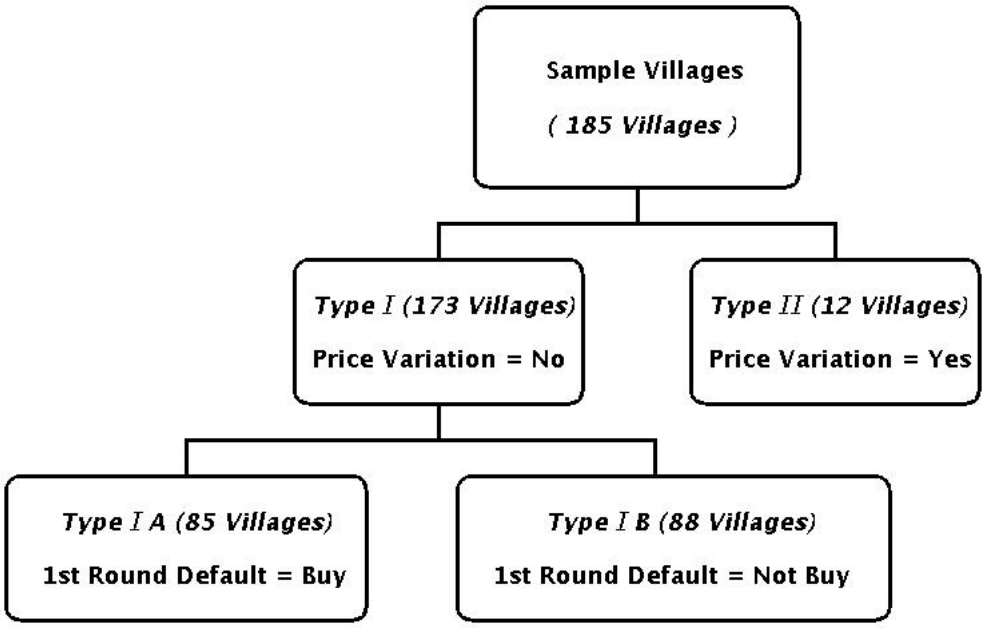


Figure 1.1. Experimental Design in Year One: Within Village Randomization

Figure 1.2. Experimental Design in Year One: Village Level Randomization



Notes: Randomizations within T3 and T4 are only available in type I villages where there was no price randomization. No additional first-round take-up information was offered to participants in T3 and T4 in type II villages.

Figure 1.3. Average Take-up Rate in Different Sessions, Year One

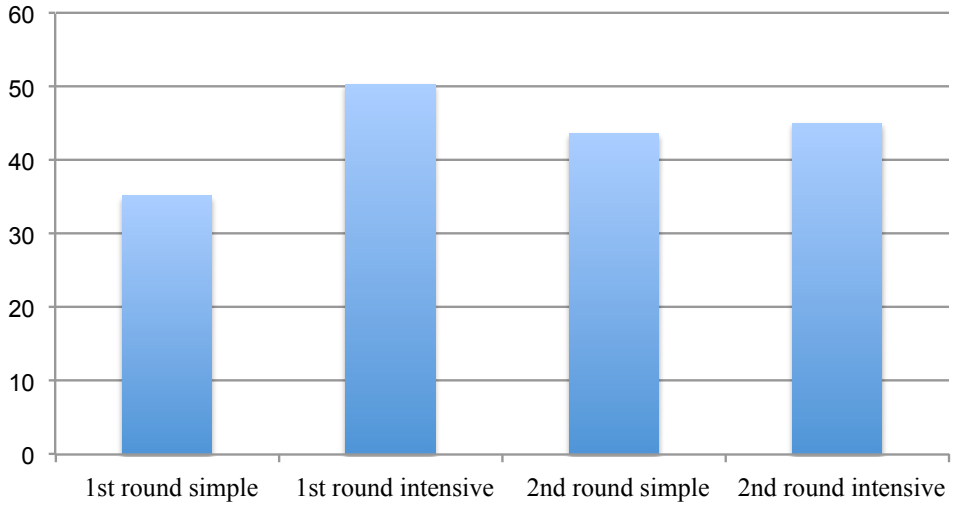
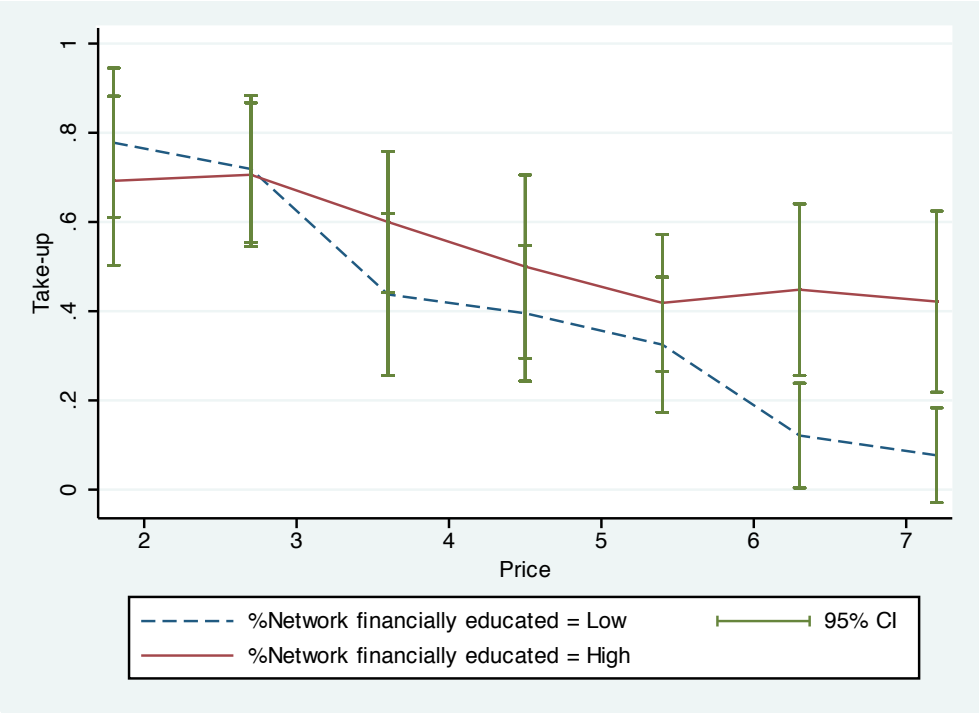
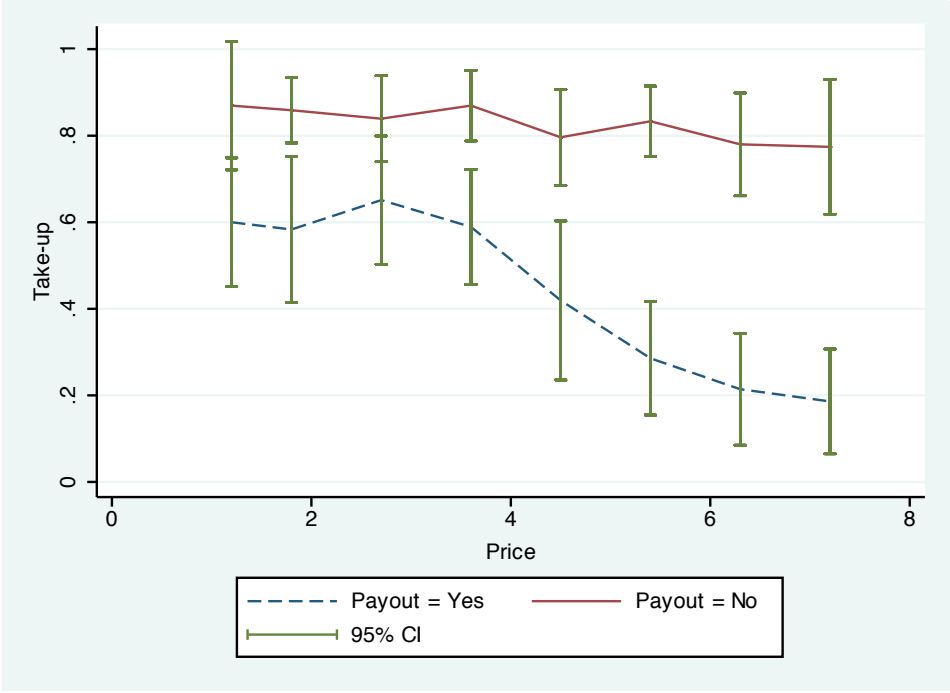


Figure 1.4. Effect of Having Friends Attending Financial Education on Insurance Demand, Year One



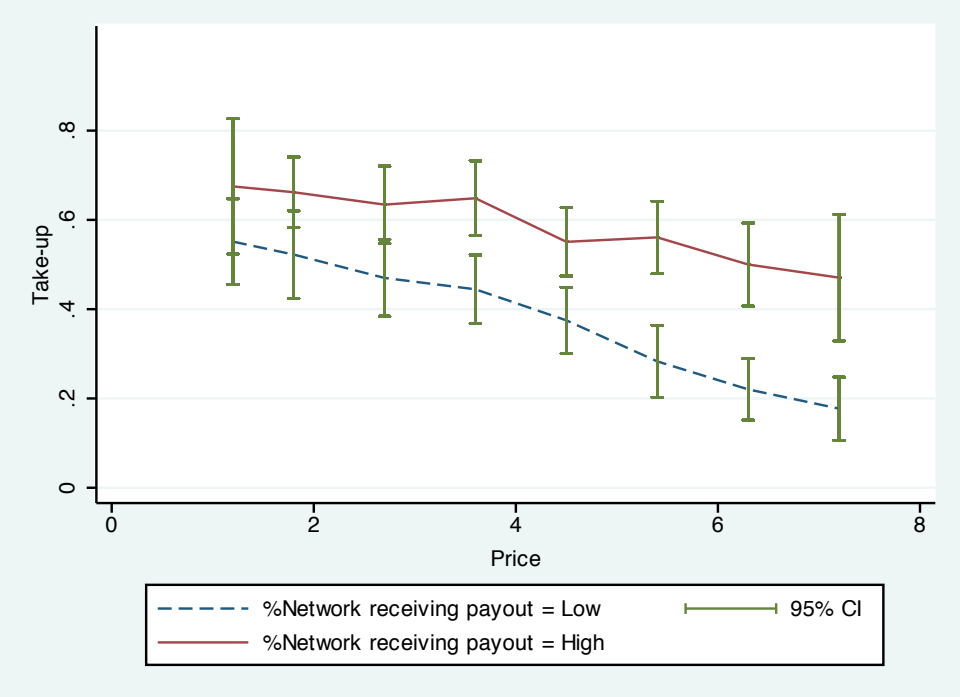
Notes: This figure is based on the sample of households in type II villages where price randomization was implemented in year one. The variable %Network financially educated is defined as "high" if it is above the sample median and is defined as "low" if it is below the sample median.

Figure 1.5. Effect of Receiving Payout on Insurance Demand, Year Two



Notes: This figure is based on the sample of households who purchased insurance in the first year.

Figure 1.6. Effect of Observing Friends Receiving Payout on Insurance Demand, Year Two



Notes: This figure is based on the sample of households who did not purchase insurance in the first year. The variable %Network receiving payout is defined as "high" if it is above the sample median and is defined as "low" if it is below the sample median.

Table 1.1 Summary Statistics

	Sample Mean	Sample Std. Dev
PANEL A: HOUSEHOLD CHARACTERISTICS		
Gender of Household Head (1 = Male, 0 = Female)	0.914	0.280
Age	51.494	12.032
Household Size	4.915	2.133
Education (0 = illiteracy, 1 = primary, 2 = secondary, 3 = high school, 4 = college)	1.192	0.853
Area of Rice Production (mu)	12.635	19.921
Share of Rice Income in Total Income (%)	73.258	34.841
Any Disasters Happened Last Year (1 = Yes, 0 = No)	0.631	0.483
Loss in Yield Last Year (%)	27.507	18.199
Risk Aversion (0-1, 0 as risk loving and 1 as risk averse), Year One	0.711	0.313
Perceived Probability of Future Disasters (%), Year One	33.633	16.619
Risk Aversion (0-1, 0 as risk loving and 1 as risk averse), Year Two	0.849	0.284
Perceived Probability of Future Disasters (%), Year Two	27.081	22.133
PANEL B: SOCIAL NETWORK MEASURES		
Number of Friends Listed	4.893	0.510
General Measure: %Friends Attending 1st Round Financial Education	0.161	0.189
Strong Measure: %Mutually Listed Friends Attending 1st Round Financial Education	0.043	0.100
Weak Measure: %2nd order Friends Attending 1st Round Financial Education	0.154	0.114
PANEL C: SOCIAL NETWORK STRUCTURAL CHARACTERISTICS		
Village Size (Number of households)	31.962	9.839
Graph Clustering (Village level)	0.182	0.085
Fraction in Giant Component (Village level)	0.987	0.070
Transtivity (Village level)	0.213	0.067
Reciprocity (Village level)	0.168	0.054
In-Degree (Household level)	3.266	2.496
Path Length (Household level)	3.578	1.941
Eigenvector Centrality (Household level)	0.148	0.098
PANEL D: INSURANCE PAYOUT, YEAR ONE		
Payout Rate Among First Year Buyers (%)	56.710	49.580
Amount of Payout Received by First Year Buyers (RMB, per mu)	86.358	66.899
PANEL E: OUTCOME VARIABLE		
Insurance Take-up Rate (%), Year One	43.941	49.637
Insurance Take-up Rate (%), Year Two	49.920	50.013
No. of Households: 5332 in year one, 1871 in year two		
No. of Villages: 185 in year one, 72 in year two		

Table 1.2. Effect of Financial Education on Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)	
	(1)	(2)
Intensive Financial Education Session	0.149***	0.140***
(1 = Yes, 0 = No)	(0.0261)	(0.0259)
Male		0.0393
		(0.0476)
Age		0.00205*
		(0.00108)
Household Size		-0.00381
		(0.00514)
Rice Production Area (mu)		0.00161
		(0.000993)
Literacy (1 = Yes, 0 = No)		0.0823***
		(0.0269)
No. of Observation	2,175	2,137
Village Fixed Effects	Yes	Yes
R-Squared	0.121	0.129

Notes: Robust clustered standard errors in parentheses. The estimation is based on the sample of participants in the two first round sessions. Village dummies are included in all estimations.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.3. Effect of Social Networks (General Measure) On Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)		
	(1)	(2)	(3)
%Network Receiving 1st Round Financial Education	0.337***	0.348***	0.489***
	(0.0810)	(0.0779)	(0.105)
Intensive Financial Education Session		0.00643	0.0539
(1 = Yes, 0 = No)		(0.0329)	(0.0397)
%Network Receiving 1st Round Financial Education			-0.301*
*Intensive Financial Education Session			(0.162)
Male		0.0374	0.0408
		(0.0673)	(0.0672)
Age		0.00374***	0.00384***
		(0.00123)	(0.00122)
Household Size		-0.00878	-0.00901
		(0.00677)	(0.00674)
Rice Production Area (mu)		0.00323***	0.00330***
		(0.00115)	(0.00114)
Literacy (1 = Yes, 0 = No)		0.0844***	0.0841***
		(0.0320)	(0.0319)
Risk Aversion		0.119**	0.114**
(0-1, 0 as risk loving and 1 as risk averse)		(0.0494)	(0.0492)
Perceived Probability of Disaster		0.00211**	0.00208**
		(0.000819)	(0.000819)
No. of Observation	1,274	1,255	1,255
Village Fixed Effects	Yes	Yes	Yes
R-Squared	0.087	0.112	0.115

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not receive 1st round take-up information from us. Social network is measured by the fraction of the five friends that a household listed who were assigned to a first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.4. Effect of Social Networks (Strong and Weak Ties) On Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)			
	(1)	(2)	(3)	(4)
%Network Receiving 1st Round Financial Education	0.447**	0.428**		
(Strong ties, mutually listed)	(0.183)	(0.182)		
%Network Receiving 1st Round Financial Education			0.0717	0.0843
(Weak Ties, second order links)			(0.148)	(0.149)
Male		0.0376		0.0397
		(0.0671)		(0.0680)
Age		0.00379***		0.00365***
		(0.00127)		(0.00127)
Household Size		-0.00838		-0.00836
		(0.00699)		(0.00694)
Rice Production Area (mu)		0.00281**		0.00288**
		(0.00120)		(0.00118)
Literacy (1 = Yes, 0 = No)		0.0778**		0.0808**
		(0.0320)		(0.0321)
Risk Aversion		0.121**		0.119**
(0-1, 0 as risk loving and 1 as risk averse)		(0.0492)		(0.0488)
Perceived Probability of Disaster		0.00220***		0.00217***
		(0.000829)		(0.000831)
Intensive Financial Education Session		0.000230		0.00305
(1 = Yes, 0 = No)		(0.0328)		(0.0330)
No. of Observation	1,274	1,255	1,274	1,255
Village Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.077	0.101	0.073	0.097

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not receive 1st round take-up information from us. Social network is measured in two ways: the strong social network is defined as the fraction of the five friends who were mutually listed and were assigned to the first round intensive session; the weak social network is defined as the fraction of second-order friends (friends' friends) who were assigned to the first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5. Effect of Social Networks (With Different Relationships) On Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)	
	(1)	(2)
%Network Receiving 1st Round Financial Education	0.339***	0.336***
(Relationship = Neighbors)	(0.109)	(0.105)
%Network Receiving 1st Round Financial Education	0.241*	0.291**
(Relationship = Relatives)	(0.131)	(0.134)
%Network Receiving 1st Round Financial Education	0.499*	0.528**
(Relationship = Government officials)	(0.274)	(0.258)
Intensive Financial Education Session		0.00704
(1 = Yes, 0 = No)		(0.0330)
Male		0.0397
		(0.0673)
Age		0.00380***
		(0.00125)
Household Size		-0.00875
		(0.00683)
Rice Production Area (mu)		0.00331***
		(0.00114)
Literacy (1 = Yes, 0 = No)		0.0828**
		(0.0319)
Risk Aversion		0.117**
(0-1, 0 as risk loving and 1 as risk averse)		(0.0497)
Perceived Probability of Disaster		0.00211**
		(0.000819)
No. of Observation	1,274	1,255
Village Fixed Effects	Yes	Yes
R-Squared	0.087	0.113

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not receive 1st round take-up information from us. Social network is measured by the fraction of five friends that a household listed who were assigned to a first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.6. Nonlinear Effect of Social Networks On Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)	
	(1)	(2)
No. of Friends Receiving 1st Round Financial Education = 1	0.0556*	0.0616*
(1 = Yes, 0 = No)	(0.0322)	(0.0319)
No. of Friends Receiving 1st Round Financial Education = 2	0.202***	0.206***
(1 = Yes, 0 = No)	(0.0418)	(0.0398)
No. of Friends Receiving 1st Round Financial Education > 2	0.282*	0.279*
(1 = Yes, 0 = No)	(0.146)	(0.156)
Male		0.0432
		(0.0667)
Age		0.00361***
		(0.00122)
Household Size		-0.00812
		(0.00669)
Rice Production Area (mu)		0.00339***
		(0.00113)
Literacy (1 = Yes, 0 = No)		0.0835***
		(0.0317)
Risk Aversion		0.116**
(0-1, 0 as risk loving and 1 as risk averse)		(0.0498)
Perceived Probability of Disaster		0.00202**
		(0.000818)
Intensive Financial Education Session		0.00467
(1 = Yes, 0 = No)		(0.0329)
No. of Observation	1,274	1,255
Village Fixed Effects	Yes	Yes
R-Squared	0.095	0.120

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not received 1st round take-up information from us. Social network is measured by the number of five friends that a household listed who were assigned to a first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.7. Heterogeneity on the Social Network Effect: Village Level Network Characteristics, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)				
	(1)	(2)	(3)	(4)	(5)
%Network Receiving 1st Round Financial Education	0.202	0.0578	-0.727	0.252	0.257
	(0.261)	(0.166)	(0.796)	(0.252)	(0.251)
Number of Households	-0.00480**	-0.00465**	-0.00403**	-0.00424**	-0.00425**
	(0.00241)	(0.00200)	(0.00184)	(0.00213)	(0.00196)
Number of Households	0.00433				
*%Network Receiving 1st Round Financial Education	(0.00745)				
Graph Clustering		-0.490**			
		(0.234)			
Graph Clustering		1.648**			
*%Network Receiving 1st Round Financial Education		(0.799)			
Fraction in Giant Component			-0.194		
			(0.325)		
Fraction in Giant Component			1.094		
*%Network Receiving 1st Round Financial Education			(0.812)		
Transtivity				-0.286	
				(0.361)	
Transitivity				0.437	
*%Network Receiving 1st Round Financial Education				(1.116)	
Reciprocity					-0.224
					(0.401)
Reciprocity					0.562
*%Network Receiving 1st Round Financial Education					(1.402)
Intensive Financial Education Session	0.0100	0.0140	0.00913	0.00644	0.0105
(1 = Yes, 0 = No)	(0.0325)	(0.0325)	(0.0326)	(0.0322)	(0.0325)
Observations	1,255	1,255	1,255	1,274	1,255
Household Characteristics	Yes	Yes	Yes	Yes	Yes
R-Squared	0.033	0.037	0.034	0.020	0.033
P-Value of Joint-significance:					
Village Level Network Characteristics	0.0868*	0.083*	0.382	0.6639	0.8471
%Network Receiving 1st Round Financial Education	0.0002***	0.0000***	0.0001***	0.0001***	0.0002***

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not received 1st round take-up information from us. Social network is measured by the fraction of five friends that a household listed who were assigned to a first round intensive session. Definitions of village level social network characteristics are in appendix B. Household characteristics include gender, age and education of household heads, household size, rice production area, risk aversion, and perceived probability of future disasters. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.8. Heterogeneity on the Social Network Effect (Year One): Who are More Influential and Who are More Likely to be Influenced?

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)					
	(1)	(2)	(3)	(4)	(5)	(6)
%Network Receiving 1st Round Financial Education	0.316***	0.544***	0.419***	0.834***	0.165*	0.273
	(0.101)	(0.189)	(0.107)	(0.213)	(0.0997)	(0.171)
Average in-degree	0.00409	0.00209				
(friends in 1st round financial education)	(0.00635)	(0.00850)				
Average in-degree		0.0186				
*%Network Receiving 1st Round Financial Education		(0.0415)				
In-degree (own)	0.00874	0.0235***				
	(0.00566)	(0.00885)				
Indegree (own)		-0.0860**				
*%Network Receiving 1st Round Financial Education		(0.0397)				
Average Out-path Length			-0.0150	-0.000249		
(friends in 1st round financial education)			(0.0124)	(0.0177)		
Average Out-path Length				-0.0666		
*%Network Receiving 1st Round Financial Education				(0.0995)		
In-Path Length (own)			-0.0128*	-0.00530		
			(0.00729)	(0.00631)		
In-Path Length (own)				-0.0680**		
*%Network Receiving 1st Round Financial Education				(0.0284)		
Average Eigenvector Centrality					0.492***	-0.0565
(friends in 1st round financial education)					(0.157)	(0.225)
Average Eigenvector Centrality						3.232***
*%Network Receiving 1st Round Financial Education						(0.948)
Eigenvector Centrality (own)					-0.0472	0.422*
					(0.174)	(0.235)
Eigenvector Centrality (own)						-2.836***
*%Network Receiving 1st Round Financial Education						(1.016)
Intensive Financial Education Session	0.00446	0.00821	0.0107	0.00926	0.00589	0.00908
(1 = Yes, 0 = No)	(0.0331)	(0.0330)	(0.0331)	(0.0328)	(0.0329)	(0.0328)
Observations	1,255	1,255	1,255	1,255	1,255	1,255
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.105	0.111	0.108	0.116	0.110	0.125
P-Value of Joint-significance:						
%Network Receiving 1st Round Financial Education		0.01***		0.0002***		0.0003***
Network Structure (friends in 1st round financial education)		0.669		0.6202		0.0001***
Network Structure (own)		0.0302**		0.0313**		0.0222**

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not received 1st round take-up information from us. Social network is measured by the fraction of five friends that a household listed who were assigned to a first round intensive session. Definitions of social network characteristics are in appendix B. Household characteristics include gender, age and education of household heads, household size, rice production area, risk aversion, and perceived probability of future disasters. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.9. Monetary Value of the Social Network Effect on Insurance Take-up, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)		
	(1)	(2)	(3)
Price	-0.112***	-0.167***	-0.151***
	(0.0162)	(0.0273)	(0.0306)
%Network Receiving 1st Round Financial Education	0.364***	-0.199	-0.241
	(0.0979)	(0.230)	(0.243)
Price * %Network Receiving 1st Round Financial Education		0.130**	0.151**
		(0.0524)	(0.0520)
Male	-0.0739	-0.0711	-0.0740
	(0.0724)	(0.0724)	(0.0741)
Age	-0.00394	-0.00354	-0.00355
	(0.00230)	(0.00249)	(0.00233)
Household Size	-0.0241**	-0.0238**	-0.0237**
	(0.0102)	(0.00997)	(0.0102)
Rice production area (mu)	-0.00146	-0.00133	-0.00140
	(0.00316)	(0.00308)	(0.00295)
Literacy (1 = Yes, 0 = No)	-0.0361	-0.0371	-0.0461
	(0.0830)	(0.0830)	(0.0933)
Risk Aversion	0.141*	0.144**	0.136*
(0-1, 0 as risk loving and 1 as risk averse)	(0.0697)	(0.0652)	(0.0672)
%Network Assigned with Higher Prices			0.0795
			(0.101)
%Network Assigned with Lower Prices			-0.0911
			(0.0770)
Observations	429	429	429
Village Fixed Effects	Yes	Yes	Yes
R-Squared	0.239	0.249	0.260
P-value of Joint-significance: Price		0.0000***	0.0013***
%Network Receiving 1st Round Financial Education		0.0057***	0.0018***

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of type II villages where seven different prices ranging from 1.8 RMB to 7.2 RMB were randomly assigned on the household level. Village dummies are included in both estimations. *** p<0.01, ** p<0.05, * p<0.1

Table 1.10. Did Social Networks Convey Insurance Knowledge?

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)		Insurance Knowledge (0 - 1)		
	(1)	(2)	(3)	(4)	(5)
Intensive Financial Education Session	0.141***		0.314***	-0.00129	
(1 = Yes, 0 = No)	(0.0259)		(0.0120)	(0.0167)	
Second Round (1 = Yes, 0 = No)	0.0901***		0.245***		
	(0.0309)		(0.0142)		
Intensive Financial Education Session *Second Round	-0.138***		-0.323***		
	(0.0422)		(0.0200)		
%Network Receiving 1st Round Financial Education		-0.106		0.356***	0.128
		(0.167)		(0.0475)	(0.103)
%Network Receiving 1st Round Financial Education		0.621***			0.312**
*Average Network Insurance Knowledge		(0.209)			(0.122)
Male	0.0477	0.0517	0.0425**	0.0393	0.0465
	(0.0343)	(0.0671)	(0.0187)	(0.0354)	(0.0352)
Age	0.00281***	0.00341***	-0.000846*	-0.000946	-0.00111
	(0.000843)	(0.00122)	(0.000435)	(0.000786)	(0.000783)
Household Size	-0.00485	-0.00825	0.00253	0.00221	0.00248
	(0.00427)	(0.00669)	(0.00247)	(0.00429)	(0.00425)
Rice Production Area (mu)	0.00167**	0.00313***	0.000457**	-0.000533	-0.000583
	(0.000788)	(0.00114)	(0.000198)	(0.000632)	(0.000621)
Literacy (1 = Yes, 0 = No)	0.0777***	0.0785**	0.0868***	0.0852***	0.0822***
	(0.0201)	(0.0316)	(0.0119)	(0.0204)	(0.0206)
Risk Aversion	0.0934***	0.108**	0.0827***	0.0495*	0.0441
(0-1, 0 as risk loving and 1 as risk averse)	(0.0265)	(0.0496)	(0.0139)	(0.0267)	(0.0269)
Perceived Probability of Disaster	0.000773	0.00209**	0.000553**	0.00145***	0.00144***
	(0.000521)	(0.000815)	(0.000277)	(0.000440)	(0.000441)
No. of Observation	3,433	1,255	3,259	1,255	1,255
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-Squared	0.093	0.118	0.233	0.132	0.137
P-value of Joint-significance:					
Intensive Financial Education Session	0.0000***		0.0000***		
%Network Receiving 1st Round Financial Education		0.0000***			0.0000***

Notes: Robust clustered standard errors in parentheses. Estimation results in columns (1) and (3) are based on households who were assigned to first round sessions or those in second round session groups without additional information. Column (2), (4) and (5) are based on households who were invited to second round sessions but did not received any additonal take-up information. Insurance knowledge is the score that a household got in ten questions that we asked during household survey to test their understanding of insurance benefits. *** p<0.01, ** p<0.05, * p<0.1

Table 1.11. Effect of the Overall 1st Round Take-up Rate on 2nd Round Decision, Year One

	First Stage: Overall 1st Round Take-up	Insurance Take-up (1 = Yes, 0 = No)			
VARIABLES		All Sample		No Information Revealed	Revealed 1st Round Overall Take-up
		OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
Default (1 = Buy, 0 = Not Buy)	0.121***				
	(0.0326)				
1st Round Overall Take-up Rate		0.388***	0.434**	0.0164	0.427*
(Village level)		(0.0710)	(0.215)	(0.338)	(0.237)
No 1st Round Take-up Information Revealed		0.128***	0.181		
(1 = Yes, 0 = No)		(0.0405)	(0.135)		
1st Round Overall Take-up Rate		-0.307***	-0.431		
*No 1st Round Take-up Information Revealed		(0.0758)	(0.318)		
Male	0.0370	0.0395	0.0395	0.0363	0.0400
	(0.0490)	(0.0439)	(0.0439)	(0.0703)	(0.0532)
Age	0.00202*	0.00480***	0.00405***	0.00374***	0.00499***
	(0.00107)	(0.000886)	(0.000806)	(0.00123)	(0.00130)
Household Size	-0.00434	-0.00474	-0.00465	-0.00773	-0.00172
	(0.00515)	(0.00502)	(0.00508)	(0.00701)	(0.00690)
Rice Production Area (mu)	0.00159	0.00149**	0.00156**	0.00166	0.00152**
	(0.000972)	(0.000625)	(0.000651)	(0.00135)	(0.000596)
Literacy (1 = Yes, 0 = No)	0.0868***	0.0751***	0.0832***	0.0757**	0.0802**
	(0.0265)	(0.0236)	(0.0228)	(0.0328)	(0.0373)
Intensive Financial Education Session		0.00750	0.00750	0.00671	0.00750
(1 =Yes, 0 = No)		(0.0180)	(0.0176)	(0.0318)	(0.0267)
Risk Aversion		0.122***	0.133***	0.131***	0.114***
(0-1, 0 as risk loving and 1 as risk averse)		(0.0163)	(0.0162)	(0.0271)	(0.0261)
Perceived Probability of Disaster		0.00183***	0.00189***	0.00175*	0.00193**
		(0.000564)	(0.000572)	(0.000890)	(0.000805)
No. of Observation	2,137	2,674	2,709	1,296	1,378
Village Fixed Effects	No	Yes	Yes	Yes	Yes
R-Squared	0.120	0.100	0.096	0.098	0.135
P-value of Joint-significance:					
1st Round Sessions' Overall Take-up Rate		0.0000***	0.116		

Notes: Robust clustered standard errors in parentheses. Column (1) present first stage results for IV estimation. Estimations from columns (2) to (5) in this table are based onthe sample of 2nd round session participants. Columns (2) and (3) are based on the whole 2nd round sample; Column (4) is based on the sub-sample who receive no extra information in addition to the presentation; Column (5) is based on the subgroup of households to whom we desseminate the first round take-up information. In IV estimations, Default options are used as the instrumental variable for the first round take-up rate. *** p<0.01, ** p<0.05, * p<0.1

Table 1.12. Effect of Friends' Decisions in 1st Round Sessions on 2nd Round Take-up, Year One

	First Stage: Network 1st round take-up%	Insurance Take-up (1 = Yes, 0 = No)			
VARIABLES		All Sample		No Information Revealed	Revealed 1st Round Decision List
		OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
1st Round Overall Take-up Rate		0.608***	0.104	0.0711	0.460
(Village level)		(0.109)	(0.732)	(0.430)	(0.790)
1st Round Network's Take-up Rate		-0.0152	0.961**	0.0996	0.969**
		(0.0528)	(0.377)	(0.252)	(0.383)
Not Revealed 1st Round Decision List		0.256***	0.441**		
(1 = Yes, 0 = No)		(0.0554)	(0.203)		
1st Round Overall Take-up Rate		-0.544***	0.0584		
*Not Revealed 1st Round Decision List		(0.124)	(0.753)		
1st Round Network's Take-up Rate		0.0192	-0.853*		
*Not Revealed 1st Round Decision List		(0.0732)	(0.464)		
Default * % Network in 1st Round Sessions	0.2829***				
	(0.0614)				
%1st Round Network in Intensive Session	0.112***				
	(0.0372)				
Male	-0.0177	0.0554	0.0479	0.0169	0.0293
	(0.041)	(0.0569)	(0.0600)	(0.0922)	(0.0732)
Age	-0.0004	0.00483***	0.00582***	0.00466***	0.00652***
	(0.0009)	(0.00106)	(0.00129)	(0.00133)	(0.00244)
Household Size	0.0028	-0.00409	-0.0116*	-0.00991	-0.00793
	(0.005)	(0.00629)	(0.00651)	(0.00791)	(0.0108)
Rice Production Area (mu)	0.0006	0.000819	0.00302***	0.00406***	0.00178
	(0.0007)	(0.00119)	(0.00102)	(0.00126)	(0.00162)
Literacy (1 = Yes, 0 = No)	0.0077	0.0858***	0.0943***	0.0943**	0.102
	(0.0276)	(0.0307)	(0.0338)	(0.0366)	(0.0628)
Intensive Financial Education Session	-0.0007	0.0237	0.0210	0.000123	0.0556
(1 =Yes, 0 = No)	(0.02)	(0.0240)	(0.0288)	(0.0359)	(0.0511)
No. of Observation	1530	1,643	1,530	920	610
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-Squared	0.1648	0.088		0.115	
P-value of Joint-significance:					
1st Round Sessions' Overall Take-up Rate		0.0000***	0.9241		
1st Round Network's Take-up Rate		0.9572	0.0355**		

Notes: Robust clustered standard errors in parentheses. Columns (1) - (3) are based on the whole second round sample. Column (4) is based on a sub-sample of the second round participants who did not received any additional information except for the presentation, while column (5) is based on households who were provided with the decision list of 1st round session. Column (1) verifies whether variables Default * network% in 1st round and fraction of friends in the 1st round who were assigned to the intensive sessions can work as valid IVs for the 1st round take-up rate among social network. *** p<0.01, ** p<0.05, * p<0.1

Table 1.13. Effect of Friends' Take-up Decisions in Year One on Second Year Insurance Demand Curve

VARIABLES	OLS Estimation		IV Estimation			
			1st Stage:	2nd Stage:		
	Insurance Take-up (Year two, 1 = Yes, 0 = No)		%Network Take-up (Year one)	Insurance Take-up (Year two, 1 = Yes, 0 = No)		
	(1)	(2)	(3)	(4)	(5)	(6)
% Network in 1st Round Sessions * Default			0.148***			
(Year One)			(0.0346)			
%Network Receiving 1st Round Financial Education			0.241***			
(Year One)			(0.0623)			
Price	-0.0504***	-0.0539***		-0.0502***	-0.0539***	-0.00487
	(0.00832)	(0.00766)		(0.00843)	(0.00765)	(0.0295)
%Network Take-up in Year One	0.118*	0.100		0.150	0.125	0.636*
	(0.0560)	(0.0649)		(0.157)	(0.165)	(0.299)
Price * %Network Take-up in Year One						-0.135
						(0.0797)
Male		-0.0171			-0.0160	-0.0228
		(0.0852)			(0.0862)	(0.0871)
Age		0.00267*			0.00265*	0.00295*
		(0.00142)			(0.00143)	(0.00148)
Household Size		0.00992			0.00986	0.0101
		(0.00709)			(0.00708)	(0.00696)
Rice Production Area (mu)		0.00373***			0.00372***	0.00399***
		(0.000922)			(0.000935)	(0.000920)
Literacy (1 = Yes, 0 = No)		0.0726**			0.0728**	0.0706**
(1 = Yes, 0 = No)		(0.0315)			(0.0310)	(0.0282)
Insurance Knowledge		0.0828**		0.0998**	0.0796*	0.0979**
		(0.0371)		(0.0393)	(0.0389)	(0.0417)
Risk Aversion		0.124***			0.123**	0.130***
(0-1, 0 as risk loving and 1 as risk averse)		(0.0386)			(0.0389)	(0.0398)
Perceived Probability of Disaster		0.00322***			0.00322***	0.00328***
		(0.000308)			(0.000313)	(0.000320)
Observations	1,782	1,741	1,783	1,782	1,741	1,741
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.084	0.13	0.142	0.087	0.130	0.120
P-value of Joint-significance: Price and Price*Network Take-up						0.0001***
Network Take-up and Price*Network take-up						0.1526

Notes: Robust clustered standard errors in parentheses. Columns (1) ~ (2) are OLS estimation results, columns (3) ~ (6) are IV estmation results with first stage results in column (3) and second stage results in columns (4) ~ (6), using %friends in 1st round sessions * Default and %friends in first round financial education as the IV for %network take-up. Variables are year two values without special comments. *** p<0.01, ** p<0.05, * p<0.1

Table 1.14. Effect of Receiving Payouts on Second Year Insurance Demand Curve

VARIABLES	Insurance Take-up (Year two, 1 = Yes, 0 = No)				
	OLS			Regression Discontinuity	
	(1)	(2)	(3)	(4)	(5)
Price	-0.0506***	-0.0831***	-0.0793***	-0.0516***	-0.0800***
	(0.0111)	(0.0127)	(0.0124)	(0.0103)	(0.0123)
Payout (1 = Yes, 0 = No)	0.402***	0.144	0.173*	0.422***	0.204*
	(0.0350)	(0.0904)	(0.0942)	(0.0529)	(0.111)
Price * Payout		0.0618***	0.0538**		0.0543**
		(0.0171)	(0.0183)		(0.0180)
Age			0.00130	0.00149	0.00124
			(0.00158)	(0.00159)	(0.00164)
Male			-0.0211	-0.0247	-0.0164
(1 = Yes, 0 = No)			(0.0796)	(0.0852)	(0.0818)
Household Size			0.00751	0.00778	0.00761
			(0.00781)	(0.00785)	(0.00772)
Rice Production Area (mu)			0.00108*	0.00129**	0.00106**
			(0.000507)	(0.000487)	(0.000460)
Literacy (1 = Yes, 0 = No)			0.0437	0.0518	0.0445
			(0.0475)	(0.0497)	(0.0469)
%Loss in Yield_history			0.000581	0.00107	0.000965
			(0.000911)	(0.00119)	(0.00120)
%Loss in Yield				-0.000728	-0.00127
				(0.00380)	(0.00360)
%Loss in Yield (square)				-1.47e-06	3.30e-06
				(3.29e-05)	(3.12e-05)
Insurance Knowledge	0.0198	0.00309	0.0109	0.0227	0.00813
	(0.0370)	(0.0380)	(0.0436)	(0.0399)	(0.0427)
Risk Aversion	0.108*	0.108*	0.0954*	0.0952*	0.0953*
(0-1, 0 as risk loving and 1 as risk averse)	(0.0518)	(0.0502)	(0.0503)	(0.0514)	(0.0497)
Perceived Probability of Disaster	0.00195***	0.00186***	0.00150***	0.00155***	0.00151***
	(0.000435)	(0.000424)	(0.000452)	(0.000462)	(0.000454)
%Network Take-up in Year One	-0.0316	-0.0231	-0.0396	-0.0467	-0.0410
	(0.0971)	(0.0835)	(0.0731)	(0.0778)	(0.0725)
Observations	735	735	718	718	718
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-Squared	0.269	0.284	0.290	0.280	0.291
P-value of Joint-significance: Price and Price*Payout		0.0002***	0.0002***		0.0003***
Payout and Price*Payout		0.0000***	0.0000***		0.0000***

Note: Robust clustered standard errors in parentheses. Results in this table is based on the sample of households who purchased insurance in the first year. Columns (1)-(3) report OLS estimation results, and columns (4) and (5) report regression discontinuity estimation results. Variables are year two values without special comments. %Loss in yield_history is defined as the average loss rate in yield in the past three years. *** p<0.01, ** p<0.05, * p<0.1

Table 1.15. Effect of Observing Friends Receiving Payouts on Second Year Insurance Demand Curve

VARIABLES	Insurance Take-up (Year two, 1 = Yes, 0 = No)					
	All Sample		1st Year Take-up = Yes		1st Year Take-up = No	
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.0499***	-0.0660***	-0.0512***	-0.0699***	-0.0464***	-0.0686***
	(0.00815)	(0.0106)	(0.0111)	(0.00999)	(0.0115)	(0.0179)
%NetworkPayout_High	0.217***	0.0816	0.0476	-0.109	0.224***	0.0407
(= 1 if % > median, and 0 otherwise)	(0.0266)	(0.0589)	(0.0317)	(0.0793)	(0.0400)	(0.0937)
Price * %NetworkPayout_High		0.0300**		0.0368*		0.0425**
		(0.0107)		(0.0177)		(0.0179)
Age		0.00231		0.00177		0.000878
		(0.00164)		(0.00191)		(0.00193)
Male		-0.0229		-0.0310		-0.0225
(1 = Yes, 0 = No)		(0.0813)		(0.0852)		(0.109)
Household Size		0.0111		0.0108		0.00902
		(0.00721)		(0.00838)		(0.00780)
Rice Production Area (mu)		0.00330**		0.00177**		0.00184
		(0.00123)		(0.000594)		(0.00127)
Literacy (1 = Yes, 0 = No)		0.0728**		0.0533		0.0564*
		(0.0310)		(0.0542)		(0.0263)
%Loss in Yield_history		0.00207**		0.000633		0.00121
		(0.000713)		(0.000924)		(0.00132)
Insurance Knowledge	0.0868**	0.0886**	0.0321	0.0366	-0.0146	-0.0148
	(0.0342)	(0.0346)	(0.0409)	(0.0536)	(0.0744)	(0.0755)
Risk Aversion	0.138***	0.121***	0.127**	0.111*	0.0974**	0.0808**
(0-1, 0 as risk loving and 1 as risk averse)	(0.0342)	(0.0363)	(0.0560)	(0.0574)	(0.0351)	(0.0354)
Perceived Probability of Disaster	0.00304***	0.00277***	0.00190***	0.00148**	0.00262***	0.00244***
	(0.000316)	(0.000295)	(0.000409)	(0.000475)	(0.000537)	(0.000509)
%Network Take-up in Year One	0.170	0.0739	0.00186	-0.0473	0.0551	-0.0177
	(0.164)	(0.156)	(0.136)	(0.102)	(0.341)	(0.346)
Payout (1 = Yes, 0 = No)			0.403***	0.396***		
			(0.0367)	(0.0354)		
Observations	1,642	1,603	671	654	971	949
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.158	0.177	0.297	0.313	0.148	0.161
P-value of Joint-significance: Price and Price*%NetworkPayout_High		0.0001***		0.0001**		0.0032***
%NetworkPayout_high and Price*%NetworkPayout_High		0.0000***		0.0723*		0.0001***

Notes: Robust clustered standard errors in parentheses. Columns (1) ~ (2) are based on the whole sample, columns (3) ~ (4) are based on the group of households who purchased insurance in the 1st year, and columns (5) ~ (6) are based on households who did not purchased in the 1st year. Variables are year two values without special comments. %NetworkPayout_high is a measure of experience of a household's social networks. Consider the fraction of a household's friends who received payou in the total number of his friends who purchased insurance in the 1st year, if it is higher than the sample median of that vairiable, %NetworkPayout_high is defined as 1, and otherwise it is defined as 0. %Loss in yield_history is defined as the average loss rate in yield in the past three years. *** p<0.01, ** p<0.05, * p<0.1

Chapter 2

Insurance Take-up and Learning from Experience

2.1 Introduction

Poor households in rural areas are vulnerable to losses from negative weather shocks. To protect themselves from these shocks, they engage in costly ex ante risk-mitigation strategies, such as avoiding high-risk and high-return agricultural activities, high levels of precautionary saving and insufficient investment in production (Rosenzweig and Wolpin (1993)) and human capital (Jensen (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 reduce poverty is to provide formal weather insurance products. However, in many cases, such insurance products exist, but their use is not widespread[1]. In 2009, a rice insurance policy was first offered to rural households in Jiangxi Province of China. Under certain reasonable assumptions, calibration suggests that more than 70 percent of rural households should buy the weather insurance (discussed in Section 2.5). However, the baseline take-up was only around 20% in our sample. These findings suggest a puzzle: why do so few households participate when weather insurance markets are available, given the potentially large benefit?

In this paper, we apply a novel method of financial education and attempt to test the role of experience and information in influencing weather insurance take-up, using a randomized experiment in rural China. Improving farmers' understanding of insurance benefits is important[2], especially in this context, because such products are new to most farmers and

[1]For example, Giné et al. (2008) found relatively low take-up (4.6%) of a standard rainfall insurance policy among farmers in rural India in 2004. Cole et al. (2011) 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.

[2]For example, in Giné 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

large disasters are relatively uncommon[3].

We offer the 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 the insurance in 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 visit 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 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 benefit of insurance, and changes in experience of disasters and insurance benefits. We investigate each of them as below.

To test whether this effect is due to the changes in risk attitudes or perceived probability of future disasters, we elicited the subjects' risk attitudes and the perceived probability of future disasters after playing the insurance games. We then test whether playing insurance games increases risk aversion or the perceived probability of future disasters, by an amount which could generate the observed 9.6 percentage points increase in take-up. However, our results show that's not the case.

To test whether this effect is from learning the benefit of insurance, we randomly assigned households to a group in which we explained the benefit of insurance by calculating the payoff of the policy under different situations, rather than by playing insurance games. However, this treatment only increased the actual take-up by 2.7 percentage points, and the increase was not statistically significant. More importantly, playing insurance games had a larger effect than just receiving calculation, a difference which is significant at 5% level for households in the no information treatment. This suggests that learning the objective benefits of insurance cannot fully explain the increased take-up.

To test whether this effect is driven by hypothetical experience of disasters, we explore a second source of exogenous variation: the number of hypothetical disasters experienced during the game. We found that, although the total number of disasters did not increase take-up significantly, the number of disasters in last few rounds did. Specifically, experiencing one more hypothetic disaster in the last five rounds increased the actual take-up by 6.7 percentage points. In the no information treatment, both the total number of disaster and

be important to help spreading good insurance product.

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

the number of disasters in last few rounds increased take-up significantly. And experiencing one more hypothetic disaster in the last five rounds increased the actual take-up by 8.5 percentage points. This suggests that experience of recent (hypothetical) disasters 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 (Giné et al. (2008); Cole et al. (2011)), lack of understanding remains less explored as a possible explanation. We provide evidence that the lack of experience and understanding of insurance contributes to the low take-up rate of weather insurance.

Second, this paper provides a new method of financial education and shows it is effective. 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 (2009b); Lusardi and Mitchell (2007); Stango and Zinman (2009)), the experimental evidence of financial education is mixed[4]. 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-just explaining the benefits.

Our results also contribute to the literature on the effect of direct experience. Existing works have shown the effect of experience in areas including consumer behaviors (Haselhuhn et al. (2009)), financial market (Malmendier and Nagel (2011)) and charitable giving. But how can we use experience to influence individual decisions is not explored yet. This paper analyzes the effect of 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 their experiences, as even hypothetical experience has an impact on household behaviors.

Fourth, this paper provides a new perspective to the literature of laboratory experiments. Laboratory experiments provide controlled institutional contexts which are otherwise exceptionally difficult to obtain and 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 that 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 laboratory experiment itself on the 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 Goette (2007)). A difference from most laboratory experiment is that we paid all households a flat fee to eliminate confounding due to income effects[5], and it is interesting that even when there's no incentives,

[4]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. (Forthcoming); Gaurav et al. (2011); Cai (2012)).

[5]The literature on financial incentives in experiments suggests that when there is no clear standard of performance in experiments like risky choices, incentives often cause subjects to move away from social

we still observed a large treatment effect. Follow-up work will tell whether experiments with monetary incentives provide similar results.

The paper proceeds as follows. In section 2.2, we provide background information on rice insurance in China. In section 2.3, we describe the experimental design and survey data. The main empirical results are discussed in section 2.4. 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 2.5, we develop a simple model to explain the results.

2.2 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 people. 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 included in the first round pilots in Jiangxi province, which is one of China's major rice bowls[6]. 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 2.1. The full insurance premium is 12 RMB per mu per season[7]. 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 the investigation of 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 is also important to note that the post-subsidy price is below the actuarially fair price according to our calculation. The profit of the insurance company equals revenue minus

desirable behavior toward more realistic choices. 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%.

[6]"Natural village" refers to the actual villages, "administrative village" refers to a bureaucratic entity that contains several natural villages.

[7]1 RMB = 0.15 USD; 1 mu = 0.165 acre. Farmers produce two or three seasons of rice each year.

payment to households and fixed cost:

$$\pi = N * Premium - N * P * Payout - FC$$

Where P is the probability of future disasters, N is the number of households who buy insurance and the indemnity is the payment to households when there is a disaster. According to the private communication with local government officials, the actual probability of a disaster that leads to 30 percent or more loss is around 10 percent. Since N*3.6 < N*10%*60, the post-subsidy price is below fair price. This implies that, as the expected benefit of purchasing insurance is positive for farmers, it is optimal for all farmers who cultivate rice to purchase it. However, since the pre-subsidy price is higher than the fair price so the insurance company earns a profit if fixed costs are not large.

2.3 Experimental Design and Survey Data

2.3.1 Experimental Design

In 2009 and 2010, we randomly selected 16 natural villages as our experiment sites. 9 hired enumerators who are consisting government officials and primary school teachers, together with the two authors, visited each village and conducted surveys of 885 households prior to the beginning of the growing season. Randomization is conducted in the household level. There were two rounds of interviews for each household. The timeline is presented in Figure 2.2.

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[8]. 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[9].

[8]As estimated by government officials.

[9]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.

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

The details of the experiment are now discussed. We first approached the leaders of the villages and obtained a list that included the names of villagers and basic information about them[10]. Then we stratified the households by their natural villages, ages of household heads, area of rice production. In each stratum, households were randomly assigned to 8 interventions. We randomized the treatments in two dimensions: how the contract was explained to the households (4 groups) and whether the true disaster risk was revealed to the households (2 groups). Figure 2.3 summarizes our design with eight groups in round 1.

The contract was explained in the following 4 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, disaster 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 1 relatively large disaster, it is better to always buy insurance in the following 10 years."

In the Game 20% (respectably, 10%) group, the enumerators followed the same procedure as in the control group and then played the hypothetical insurance games with 20% (10%) probability of disaster for 10 rounds. The game was played in the following way. Household heads were first asked whether they would like to buy insurance in 2011 hypothetically and then played a lottery with 20% (10%) probability that there is 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 9 rounds from hypothetical year 2012 to year 2020[11]. In the end of the game, we gave households the same information as in the Calculation group. Note that the game treatment provides not only financial education but also the second source of randomization: the number of the hypothetical disasters experienced during the games is randomized.

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

[11]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.

In a crossed randomization, we also randomized whether households were informed of the actual probability of disaster at the end of round 1, which is 10% according to the private communication with local government officials. This randomization is interacted with how the contract is explained and 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. The comparison of these measures between the game group and other groups allows us to test whether playing games changes these parameters and further change the actual insurance take-up[12]. 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 Table B.1. 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 also use a simple mechanism to illustrate probability which might be difficult to think for rural households[13].

2.3.2 Survey Data

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

Table 2.1 presents summary statistics and balance check separately for each wave. In total, we visited 885 households in round 1 and 816 households in round 2. The overall attrition rate between round 1 and round 2 is 7.8 percent. While the attrition is slightly higher in the game group, 9.8 percent, than in the control and calculations groups, 6.2 and

[12]We did not ask them the same questions before the games because households might be consistent with themselves so that we could not see the treatment effects of these two measures.

[13]The enumerators gave sample individuals 10 small paper balls and asked them to put these paper balls into two areas: (1) no disaster more than 30% next year and (2) disaster more than 30% next year. If households put 2 paper balls into (2) and 8 paper balls into (1), their perceived probability of future disaster is around 20%.

5.6 percent, the difference of attrition between groups is not statistically significant. The attrition is 11.8 percent in the information group, which is not significantly different from that in the no information group in wave 3, 10.4 percent.

The summary statistics shows 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.

2.4 Empirical Result

2.4.1 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 2.4, 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. So both the game and the calculation treatment increase take-up but the game treatment is more effective.

Figure 2.5 shows the treatment effects of the game treatment and the calculation treatment by the information treatment. In the no information group, the pattern is similar to Figure 2.4. However, the game treatment increases the take-up and it 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.

We estimate the treatment effect on the take-up decision through a logit regression in (2.1):

$$buy_{ij} = \alpha_j + \alpha_k + \beta_g Tg_{ij} + \beta_c Tc_{ij} + \phi X_{ij} + \epsilon_{ij} \tag{2.1}$$

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 β_g is an unbiased estimate of the reduced-form intention-to-treat (ITT) game treatment effect and β_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 α_j and α_k are village fixed effects and enumerator fixed effects, respectively. ϵ_{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 2.2.

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, the

village and enumerators fixed effects. The marginal effect of the game treatment (0.096) is positive and significant at the 5% level. So the game treatment increases the take-up by 9.6 percentage points and it is around 48 percent increase relative to 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, which are negative and not statistically significant. The difference of 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 sample. The pattern is similar to column 1. The marginal effect of the percentage of loss last year is 0.22% and it is significant at the 10% level. Therefore, 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 sample. The pattern is still similar to column 1. Education level is positively correlated with the 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 of actual loss during the previous year.

2.4.2 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.

Risk Attitudes

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

$$buy_{ij} = \alpha_{2j} + \beta_{risk} risk_{ij} + \beta_{prob} prob_{ij} + \delta_{ij} \tag{2.2}$$

$$risk_{ij} = \alpha_{3j} + \gamma_{gr} Tg_{ij} + \gamma_{cr} Tc_{ij} + \eta_{ij} \tag{2.3}$$

$$risk_{ij} = \alpha_{4j} + \beta_{dr} disaster_{ij} + \omega_{ij} \tag{2.4}$$

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 (2.2) analyzes the correlation between the take-up and risk attitudes. We restrict the sample to the control group and the calculation group in Equation (2.2) because we asked them questions about their risk attitudes and the perceived probability of future disasters before any intervention took place. In Equation (2.3) and (2.4), we estimate the effects of playing games and experiencing disasters, respectively. We assume that there is no measurement error of risk attitudes and perceived probability of future disaster, and the estimation in Equation (2.2) is unbiased.

In column 1 of Table 2.3, estimates from (2.2) are presented. The coefficient t 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 (2.3), 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 (2.4). 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 (2.1), (2.2), and (2.3), generate indicators for each equation and estimated the regression system. To account for the correlation of error term between each equation, standard errors are clustered by 93 natural villages[14]. 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 (2.1), (2.2), and (2.4) and estimate the regression system. We test the hypothesis: $1.48\beta_{risk}\gamma_{gr} = \beta_g$, where 1.48 is average number of hypothetical disasters they experience during the games. We reject the hypothesis at the 5% level (p=0.044), with the 95% confidence interval of $1.48\beta_{risk}\gamma_{gr}$ 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.

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:

$$prob_{ij} = \alpha_{3j} + \gamma_{gp} Tg_{ij} + \gamma_{cp} Tc_{ij} + \eta_{ij} \tag{2.5}$$

$$prob_{ij} = \alpha_{4j} + \beta_{gp} disaster_{ij} + \omega_{ij} \tag{2.6}$$

[14]The procedure is discussed in Appendix B.

where $prob_{ij}$ is the perceived probability of future disaster. In Equation (2.5) and (2.6), we estimate the effects of playing games and experiencing disasters, respectively. The results of (2.5) and (2.6) are presented in column 4 and 5 in Table 2.3, respectively.

The treatment has negative effect on the perceived probability of future disasters in columns 4. The coefficient of the number of hypothetical disasters is not significant. Following the similar procedures in the last section, we test the hypothesis $\beta_{prob}\gamma_{gp} = \beta_g$ and $1.48\beta_{dp}\gamma_{gp} = \beta_g$ to determine whether changes of 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 the similar procedures in the last section 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.

Learning the Benefit 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 because the difference between those two interventions should indicate whether households acquire disaster experiences during the games.

We run various regressions with (2.1) and report the p-value of Ward test $\beta_g = \beta_c$ in Table 2.2. In columns 1, 4 and 5, we use the whole sample. The difference between β_g and β_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 β_g and β_c is around 11 percentage points and it is significant at the 5% level.

In sum, when we consider the channel of the game treatment effect without interaction effect of the information treatment, learning the benefit 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, it is only suggestive evidence that learning the benefit is unlikely to explain the game treatment effect.

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 2.6 about actual take-up and the hypothetical disasters experienced during the games.

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%, the take-up of households who experienced one disaster is higher than those who experienced zero or two and more disasters. But given the relatively large standard deviation, Figure 2.6 provides only suggestive evidence that the take-up rate is increasing in the number of hypothetical disasters experienced and 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:

$$buy_{ij} = \alpha_j + \beta_{disaster} disaster_{ij} + \delta_{ij} \tag{2.7}$$

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

The marginal effect estimated in (2.7) is presented in column 1 and 4 of Table 2.4. 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 (2.8) to analyze these two effects:

$$buy_{ij} = \alpha_j + \beta_0 disaster0_{ij} + \beta_1 disaster1_{ij} + \beta_2 disaster2_{ij} + \beta_3 disaster3_{ij} + \epsilon_{ij} \tag{2.8}$$

where $disasterK_{ij}$ is an indicator that takes on a value of one if households experience K disasters during the games. β_0 captures the understanding effect and the difference between β_0 and other coefficients captures the vividness effect.

The marginal effect of (2.8) is presented in column 2 and 5 in Table 2.4. 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 as we have seen in Table 2.2. The difference between β_1 and β_2 is statistically significant at the 10% level. However, we cannot reject the hypothesis that β_0 and β_1 are the same. Therefore, we could not distinguish the understanding effect and 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 last 5 rounds in Figure 2.7. The evidence in Figure 2.7 suggests that the number of hypothetical disasters experienced in the first 5 rounds cannot influence take-up, whereas that 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:

$$buy_{ij} = \alpha_j + \beta_{f5} disasterfirst5_{ij} + \beta_{15} disasterlast5_{ij} + \delta_{ij} \tag{2.9}$$

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[15]. 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[16]. These results are consistent with the literature in experienced utility and recency effects (Fredrickson and Kahneman (1993); Schreiber and Kahneman (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 disaster and the number of disasters in last few rounds increase take-up significantly. These results suggest that experience of recent hypothetical disasters might be the channel through which the games influence insurance decisions.

2.5 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 effect could occur. In section 2.5.1, we show that standard constant absolute risk aversion (CARA) preference and constant relative risk aversion (CRRA) preference is unlikely to explain the data. In section 2.5.2, we add a weight parameter in the utility function to capture the influence of experience. Then we estimate the parameters through a maximum likelihood method (MLE).

2.5.1 Standard Model

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

$$u(x) = -\frac{exp(-\alpha x)}{\alpha} \tag{2.10}$$

With CARA preferences, the consumer's wealth does not affect his insurance choices. Therefore, the take-up decisions should be determined by the joint distribution of risk attitudes and the perceived probability of future disasters.

Let $U(a)$ denote the household utility as a function of insurance decision. $a = 1$ if the household buys the insurance and $a = 0$ if the household does not buy the insurance. Let (b, τ) denote the insurance contract in which b is the repayment of insurance if there is a

[15]See Appendix Table B.4 for detail.

[16]See Table B.4 for detail.

disaster and τ is the premium. Let x to be the gross income of rice production, p to be the perceived probability of future disasters and l denote the loss in yield. The expected utility of not buying the insurance is:

$$U(a=0) = (1-p)u(x) + pu(x-l) \tag{2.11}$$

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

$$U(a=1) = (1-p)u(x-\tau) + pu(x-\tau-l+b) \tag{2.12}$$

The condition for the household to buy the insurance is:

$$U(a=1) \geq U(a=0) \tag{2.13}$$

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 preference 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 (α) and perceived probability of future disasters (p). Although we do not observe parameter α, we can make use of their choices in Table 2.1 to estimate the intervals of their α in the utility function. The intervals of α under CARA and CRRA are presented in Table 2.5. If a household takes two riskless options, α should be greater than zero and less than 0.0041 under CARA preference. The details of the simulation procedures are discussed in Appendix B.

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 preference is unlikely to explain our data.

Another route is to ignore α and p as elicited. Suppose that we had not elicited measures for risk aversion and perceived probability of future disasters. Then we estimate α and p in the logit formula (2.14) through MLE:

$$P(a=1) = \frac{exp(U(a=0))}{exp(U(a=1)) + exp(U(a=0))} \tag{2.14}$$

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

In sum, both the calibrated decisions and estimated parameters contradict our data under standard CARA and CRRA preference. These results suggest that standard CARA and CRRA preference is unlikely to explain the increased take-up and the observed probability of future disasters together.

2.5.2 Model Based on Experience

We have shown that standard CARA and CRRA preference is unlikely to explain the data. In order to develop a model that fit 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 experienced the hypothetical disasters during the games. We develop a simple model in the following.

$$U(a=0) = (1-p)u(x) + pu(x-\mu l) \tag{2.15}$$

$$U(a=1) = (1-p)u(x-\tau) + pu(x-\tau-\mu l+\mu b) \tag{2.16}$$

where μ is a parameter that measures the weight to disaster loss and insurance benefit. The idea is that households might give less weight to disasters and benefits which they experienced 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 benefit and increase the weight parameter μ.

It is straightforward to show that under the assumption of CARA preference with inattention parameter μ, if $\alpha > 0$, $\frac{\partial P(buy==1)}{\partial \mu} > 0$. To the extent that playing games increases μ, it would increase the insurance take-up. To test this, we allow μ in the group who do not play games (μ_1) to be different from μ in the group who play games (μ_2). Then we estimate μ_1 and μ_2 with MLE and simulation. The details of the estimation procedures are discussed in Appendix B.

The result is presented in column 2 table 2.6. We find the estimated mean of μ_1 is around 0.21 and μ_2 is around 0.37. 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 preference. Although the point estimates are different, the key pattern is similar. These results are consistent with our hypothesis that playing games increase μ and thus increase the insurance take-up.

Hypothetical experience might have two effects: changes in understanding and changes in vividness. We add another parameter δ in (2.16):

$$U(a=1) = (1-p)u(x-\tau) + pu(x-\tau-\mu l+\mu b) + \delta \tag{2.17}$$

where δ 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 that they understand. It might capture ambiguity aversion and it is a reduced form

in our model and. We normalize δ to be zero in the game treatment so that the estimated δ is the difference of the utility of understanding. We estimate μ_1, μ_2 and δ with the same procedure to estimate μ_1 and μ_2 and present the result in column 3.

The estimated mean of μ_1 is around 20.4% and μ_2 is around 33.9%. T-test and Kolmogorov-Smirnov test show that the mean and the distribution are significantly different at the 1% level. The estimated mean of δ is -1.097 and t-test shows that the mean is significantly different from zero at the 1% level. Since we normalize δ to be zero in the game treatment, this means 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 increase the understanding and vividness and thus increase the insurance take-up.

In order to understand the empirical relationship between experience and the weight parameter, we model μ in the following way:

$$\mu = C + D(1 - exp(-K_a f_a - K_h f_h) \tag{2.18}$$

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

f_a is actual experience and it is measured by percentage of disasters more than 30% in the past 3, 2 or 1 year. f_h is hypothetical experience and it is 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, k_h and compare the effect of actual and hypothetical experience.

In column 4, we estimate both the learning effect of actual experience (k_a) and hypothetical experience (k_h) under CARA preference. f_a is measured by percentage of disasters 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 and they are significantly different at the 1% level. Column 8 presents the result with CRRA preference. Although the point estimates are different, the key pattern is similar. These results suggest that both actual and hypothetical experience matter. Moreover, hypothetical experience a few days prior to the insurance decision have stronger effects than actual experiences one year prior.

2.6 Conclusions

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 increased the actual insurance take-up by 9.6 percentage points, a 48% increase relative to the baseline take-up of 20 percentage points.

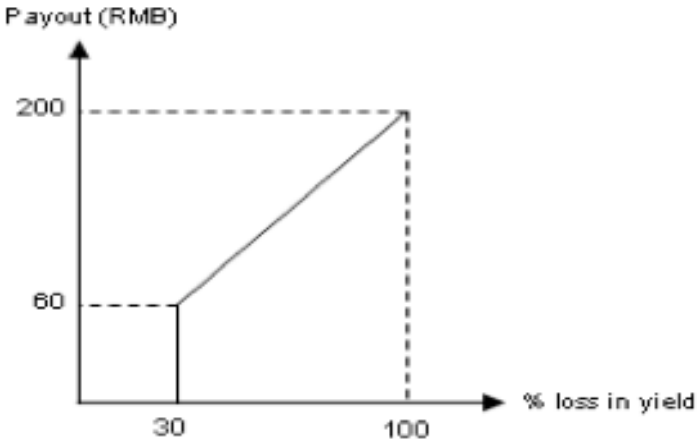
We investigate the possible mechanisms through which this effect could work, and find that changes in experience of disasters and insurance benefits are very likely to be the channel.

Whether financial education is effective to change individual decisions? The evidence in the literature is mixed. Why is financial education effective sometimes but not effective in other times? Under what circumstance s is financial education effective? This paper shows that financial education with simulated experiences can help to 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 of individual participation and then apply specific financial education to remove the barriers. This seems to work better than general financial education.

From policy perspective, this paper suggests that just using subsidy to spread new financial products might not work effectively. Our calibration (in Section 2.5) shows that, although the insurance is highly subsidized, rural households still buy less insurance than they should when they voluntarily make their decisions. This has two policy implications. First, policy makers should take into account the individual bias when design policies, especially in rural areas where most people are less educated. Second, when price tools do not work, policy makers might design a nudge (Thaler and Sunstein (2003), Thaler and Sunstein (2008)), such as defaults, to push individual decisions to the right direction.

From methodological standpoint, this paper is among the first to use a laboratory experiment as an intervention in the field experiment. We find that the laboratory experiment influences the field behavior in our setting. By using laboratory experiment, we can explicitly manipulate more variables which are endogenous or difficult to manipulate before. For example, Malmendier and Nagel (2011) find that individuals who have experienced low stock-market returns throughout their lives so far are less likely to participate in the stock market. But it is difficult to manipulate experience in order to influence individual decisions. In this paper, we use a laboratory experiment to simulate their experiences and influence their field behaviors. We hope to explore in future research whether this can apply to other settings.

Figure 2.1. Insurance Contract



Notes: 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.

Figure 2.2. Timeline

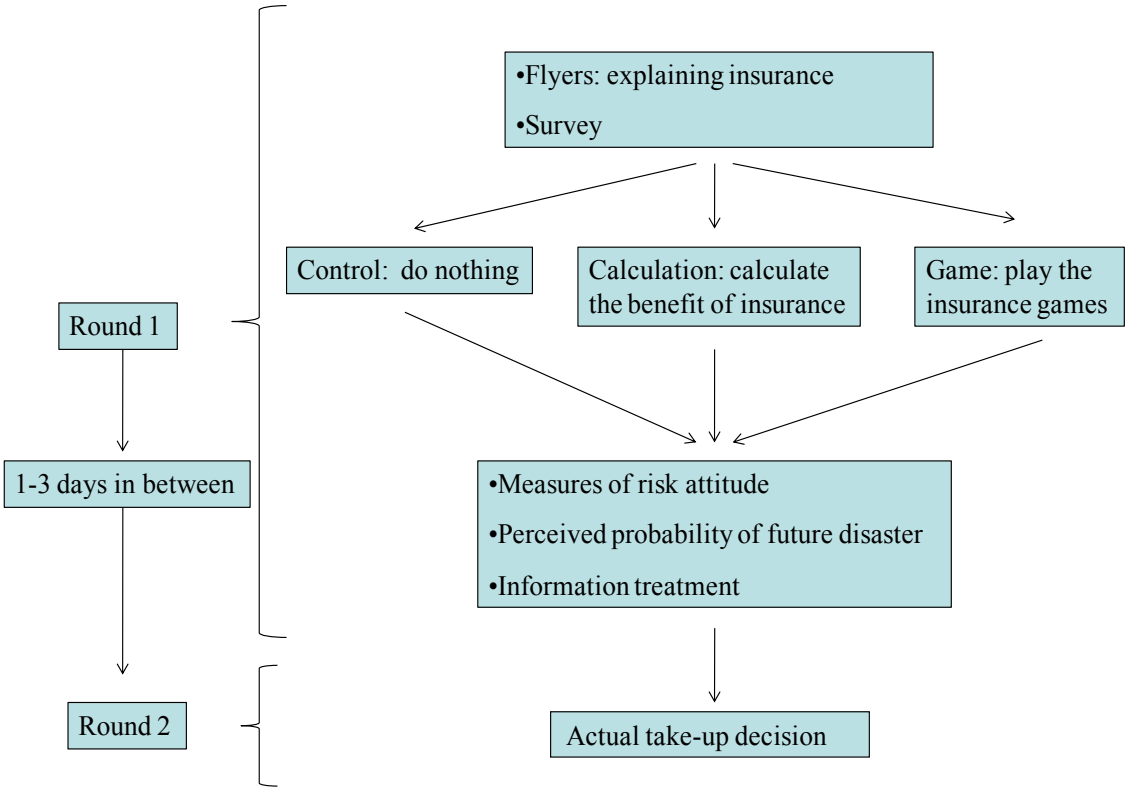


Figure 2.3. Interventions

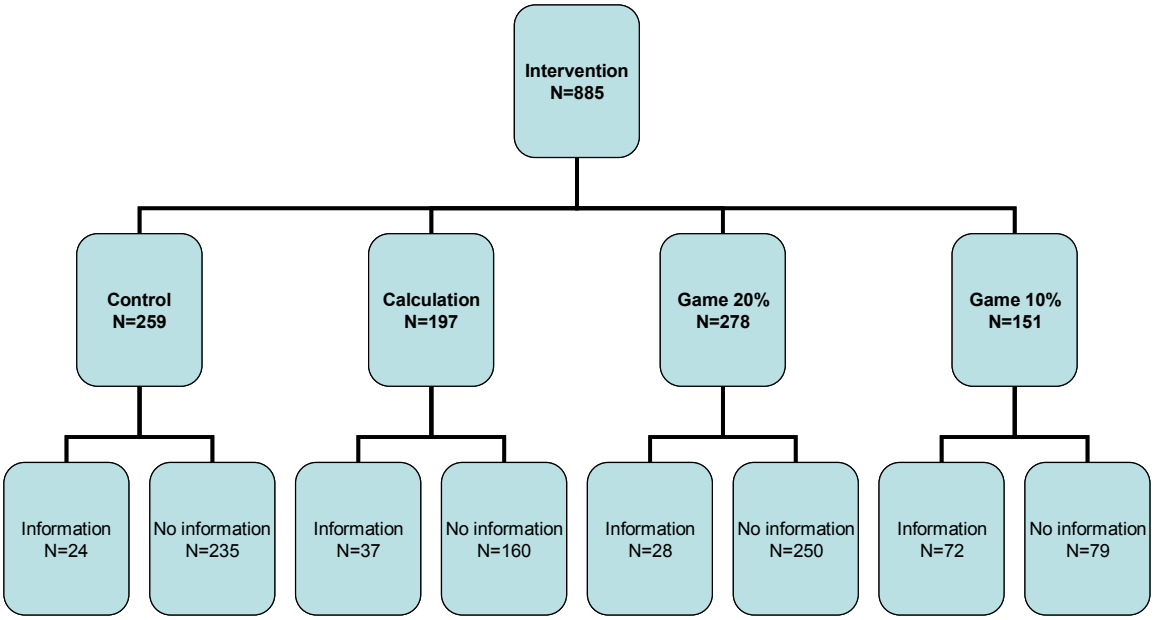
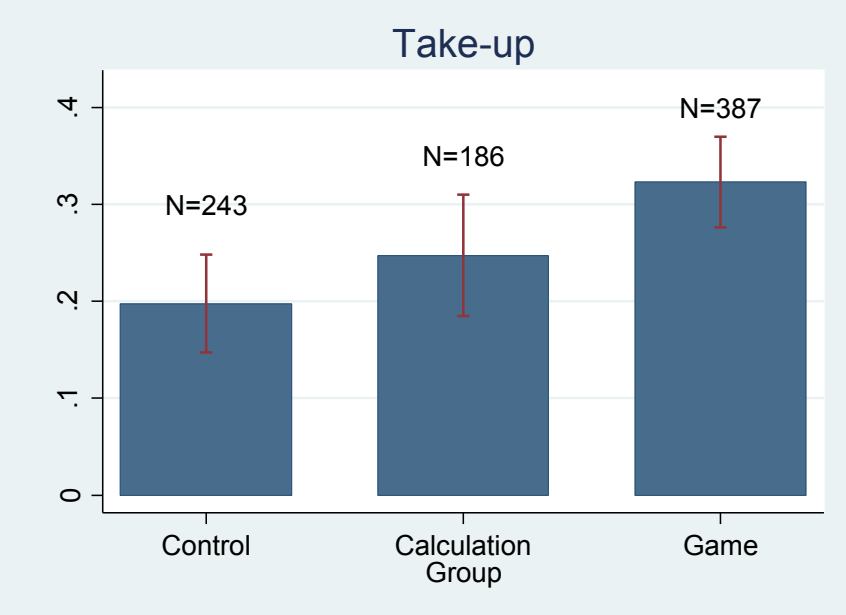
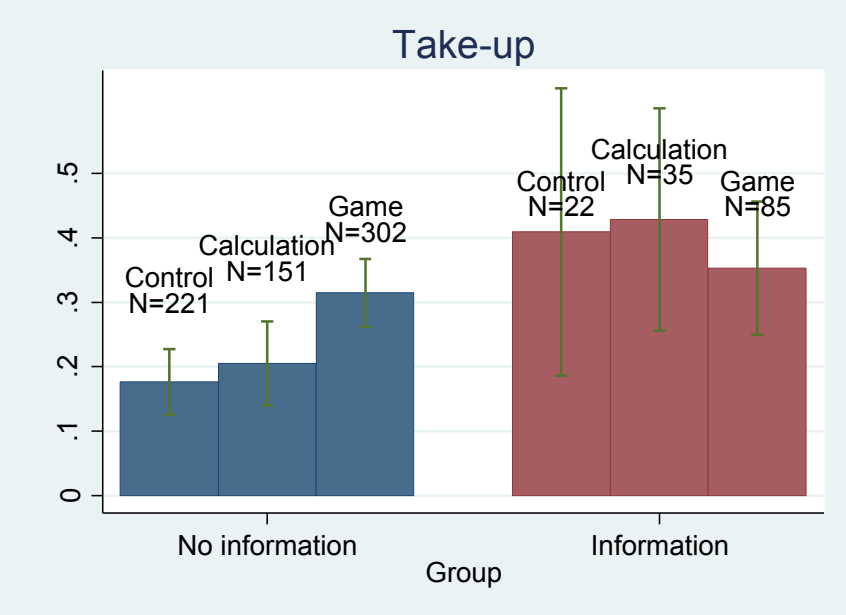


Figure 2.4. Treatment Effect



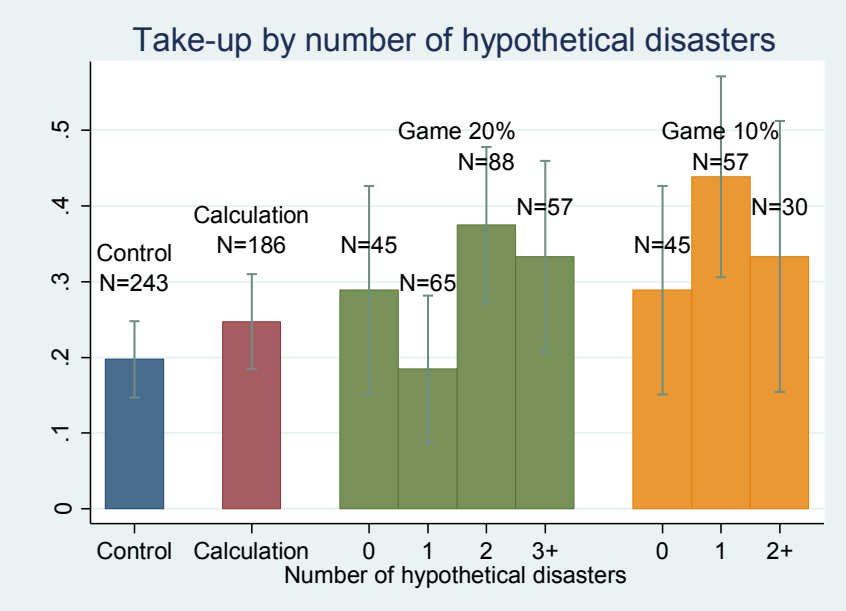
Notes: This figure shows the treatment effect of the calculation group and the game group. In the control group, the take-up is 19.8%. In the calculation group, the take-up increases to 24.7%.In the game group, the take-up increases to 32.3%. It suggests that both the game treatment and the calculation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

Figure 2.5. Treatment Effect by the Information Treatment



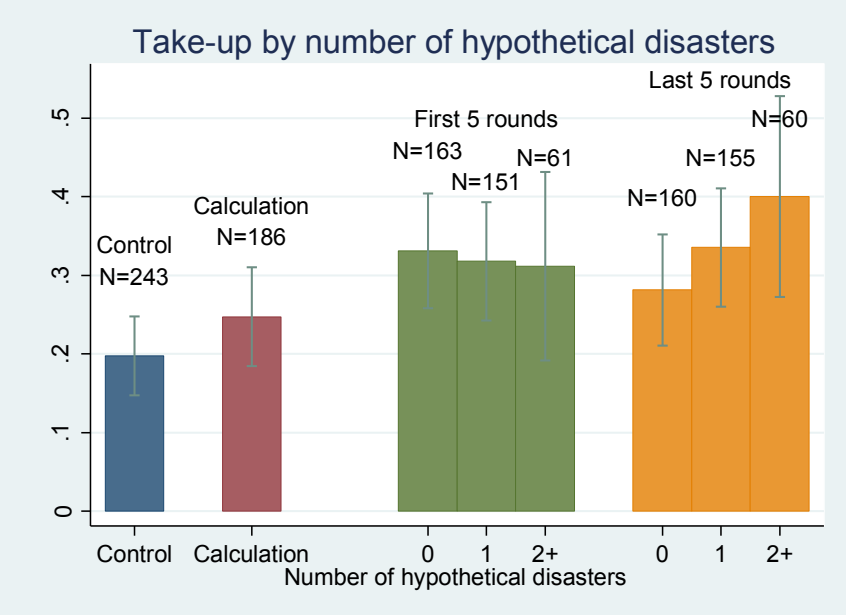
Notes: This figure shows the treatment effect by the information treatment. Without the information treatment, the game treatment is more effective than the calculation treatment. With the information treatment, the game treatment and the calculation treatment is not effective.

Figure 2.6. Take-up by Number of Hypothetical Disasters in the Games



Notes: 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.

Figure 2.7. Take-up by Number of Hypothetical Disasters in Different Rounds



Notes: 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.

Table 2.1. Summary Statitics 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	

Notes: Standard deviations are in the parentheses. *P-value in wave 1 is for F test of equal means of two groups. **P-value in wave 2 and 3 are for Wald test of equal means of three and four groups. ***Education is coded as follows: 0-illiteracy; 1-primary school; 2-secondary school; 3-high school; 4-college

Table 2.2. 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.

Table 2.3. 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.

Table 2.4. 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.

Table 2.5. 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	$\alpha<-0.0121$	$\alpha<-1.4$
1	$-0.0121<\alpha<-0.0041$	$-1.4<\alpha<-0.35$
2	$-0.0041<\alpha<0$	$-0.35<\alpha<0$
3	$0<\alpha<0.0041$	$0<\alpha<0.25$
4	$0.0041<\alpha<0.0121$	$0.25<\alpha<0.5$
5	$\alpha>0.0121$	$\alpha>0.5$

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

Table 2.6. Maximum Likelihood Estimation of Utility Function

	CARA			CRRA		
	(1)	(2)	(3)	(4)	(5)	(6)
μ_1	0.208	0.204		0.152	0.149	
μ_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 μ_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 μ_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 and CRRA utility function through MLE. In all columns, we constrain a 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 a . In column 1 and 4, we allow the weight parameter in the group who do not play games (μ1) to be different from weight parameter in the group who play games (μ2). In column 2 and 5, we add δ to measure the utility of understanding the insurance if they buy the insurance. We normalize δ to be zero in the game treatment so that the estimated δ is the difference of the utility of understanding. In column 3 and 6, we assume that the weight parameter has the following structure μ = C+D(1-exp(-kafa-khfh)) . Then we estimate both the learning effect of actual experience (ka) and hypothetical experience (kh) with different measurement of actual disaster.

Chapter 3

The Impact of Insurance Provision

3.1 Introduction

Poor households in rural areas are exposed to substantial negative shocks such as weather disasters, which can generate large fluctuations in income and consumption if insurance markets are incomplete. To protect themselves from these risks, rural households undertake costly risk management and coping strategies such as informal insurance, avoiding high risk-high return agricultural activities, holding precautionary savings, and reducing investment in production, etc (Morduch (1995), Rosenzweig and Stark (1989)). However, existing evidence shows that informal insurance mechanisms cannot effectively reduce negative impacts of regional weather shocks (Townsend (1994)). In the absence of formal insurance markets, the negative shocks and foregone profitable opportunities can lead to highly variable household income and persistent poverty (Dercon and Christiaensen (2011), Jensen (2000), Rosenzweig and Wolpin (1993)).

Although many developing countries have started to develop and market formal insurance products to shield farmers from risks, take-up is usually surprisingly low even with heavy government subsidies[1]. While there is a growing literature studying ways to improve insurance demand (Cole et al. (2011), Cai (2012), Cai and Song (2011), Bryan (2010)), rigorous evaluations of the impacts of insurance provision are quite rare. With a rich household level panel data (2000-2008) from the Rural Credit Cooperative (RCC)[2] of China, this paper studies the effect of insurance provision on household's production, borrowing, and saving decisions. The program I am studying is a weather insurance policy for tobacco farmers offered by the People's Insurance Company of China (PICC) starting from 2003 in one county of Jiangxi province. It was expanded to more areas afterwards and was implemented province-wide at

[1]For example, Giné et al. (2008) found a low take-up (4.6%) of a rainfall insurance policy among farmers in rural India in 2004, while Cole et al. (2011) found an adoption rate of 5% - 10% of a similar insurance policy in two regions of India in 2006.

[2]RCC is the most important financial institution in rural China. It is the main provider of microcredit, and most farmers have saving accounts there.

the beginning of 2010. Purchase of insurance was made compulsory for tobacco farmers in treatment regions. I take advantage of the variation in insurance provision across both regions and household types (tobacco households v.s. other households) to estimate the effect of insurance provision on household behavior, focusing on the initial stage of the policy in 2003.

The empirical strategy includes both difference-in-difference (DD) and triple difference (DDD) estimations. I use tobacco households outside of the treatment region to control for industry-specific trend in outcomes, and use non-tobacco households both within and outside the treatment region to control for region-specific trend in the absence of the policy intervention. Then the extra changes in household behavior for tobacco households in treatment regions can be attributed to the insurance policy implementation. I find that first, insurance provision has a statistically significant positive effect on the production of insured crop: it raises tobacco production by around 22% and decreases production diversification by around 29%; second, insured households tend to borrow more from the rural bank for investment on tobacco production, and the magnitude of effect is about 25%; third, the insurance policy decreases household saving rate by more than 30%; forth, estimation of dynamic effects shows that while the effect of insurance policy on both borrowing and saving became significant shortly after the policy was implemented, the impact on borrowing is persistent through the end of the sample period, while the effect on saving became significant several years after the intervention and decreased towards the end of the sample period.

This paper contributes to the existing literature in the following ways. First, it provides insights on the insurance take-up and impact literature. Estimating the causal effect of insurance policy on household behavior is made challenging by the endogenous insurance purchase decisions. There are a few papers studying effects of insurance markets on household investment, welfare, and loan repayment, using different estimation strategies. For example, Cole et al. (2011) use a randomized experiment which provided free rainfall insurance for a selected group of landowner farmers in a semi-arid area of India, and find that provision of insurance induced farmers to shift production towards higher-return but higher-risk cash crops. Gin and Yang (2009) implemented a field experiment in Malawi which offered crop technology loans to 800 maize producers and randomly bundled insurance with loans for half of them. They found the surprising result that the take-up of loans was 13% lower for producers who were offered insurance with a loan. Carter et al. (2007) analyze two index-based insurance schemes in Peru and used simulations to show that insurance provision did significantly improve producers' welfare, credit supply, and loan repayment. In contrast, Rosenzweig and Wolpin (1993) show by simulation that the gain from weather insurance for Indian farmers was minimal due to the existence of informal insurance mechanisms. This paper complements the existing literature by testing the effect of insurance provision on households' saving behavior, taking advantage of the administrative saving data from the rural bank; Since purchase of insurance in treatment regions were compulsory, take-up decisions are not endogenous here. Moreover, it employs a rigorous estimation strategy to remove confounding factors, and to estimate not only short-term effects, but also long-term

impacts. In addition, since large and significant impacts of insurance policy are found in this paper, it supports the proposition that studying ways to improve voluntary insurance take-up is important.

Second, the paper contributes to the literature explaining low investment and technology adoption in developing countries. Credit constraints and the lack of information or knowledge are often proposed as explanations (Feder et al. (1985)). Duflo et al. (2011) argue that behavioral biases limit profitable agricultural investments. This paper shows that riskiness of these investments is an important barrier, and reducing the riskiness can improve investments persistently.

The rest of the paper is organized as follows. Section 3.2 describes the background for the study and the insurance contract. Section 3.3 explains the data and summary statistics. Section 3.4 presents estimation strategies and results, and section 3.5 concludes.

3.2 Background

Tobacco is an important cash crop in China. There are more than 2,000,000 rural households who live on tobacco production. The net profit of tobacco production is around 2000 RMB per mu[3], which is 3 to 5 times that of food crops like rice.

In China, most tobacco producing counties are poor and mountainous areas. In the province that I study, there are 12 main tobacco production counties. Those counties are in two agricultural cities[4], Fuzhou and Ganzhou. Nearly half of those 12 counties are national poverty-stricken counties. To improve poverty, from the late 1990s, these counties started to develop highly profitable tobacco industries by encouraging farmers to cultivate tobacco, organizing tobacco associations to teach farmers production techniques, etc.. Taxes on tobacco production is now the main source of government revenue in these counties.

However, as other crops, tobacco production can be greatly influenced by weather risks. For example, in 2002, a flood happened and destroyed most tobacco production in some of those 12 counties, which caused huge losses in household income and government revenue. Because the vice-head of Guangchang county who is charge of fiscal and finance was previously a manager of an insurance company, he started to think about cooperating with insurance company to shield tobacco farmers from frequent weather disasters in order to give them more incentives to continue tobacco production. In 2003, the People's Insurance Company of China (PICC) designed and offered the first tobacco production insurance program to households in Guangchang county. The policy was extended to some other counties afterwards.

The insurance contract is as follows. The actuarially fair price estimated by the insurance company is 12 RMB per mu. The county and town level government gives a 50% subsidy on the premium, so farmers only pay the remaining half of around 6 RMB per mu. All tobacco

[3]1 RMB = 0.15 USD; 1 mu = 0.067 hectare.

[4]In China, city is a larger geographical unit than county.

households whose main source of income is tobacco production were required to buy the insurance for all their tobacco areas. The insurance covers natural disasters including heavy rain, flood, windstorm, extremely high or low temperature, and drought. If any of the above natural disasters happened and led to a 30% or more loss in yield, farmers were eligible to receive payouts from the insurance company. The amount of payout increases linearly with the loss rate in yield, with a maximum payout of 420 RMB. The loss rate in yield is investigated and determined by a group of insurance agents and agricultural experts[5]. The average net income from cultivating tobacco is around 2000 RMB per mu, and the production cost is around 400 RMB to 600 RMB per mu (not including labor cost). Thus, this insurance program provides partial insurance that covers around 20% of the gross income or most of the production cost.

3.3 Theoretical Model

Here I am providing a two-period, two state model to show how the provision of insurance influences farmers' investment and financial decisions[6]. Intuitively, in the first period, insurance provision increases farmers' investment on production because the expected income from production is higher in that case. As a result, it has a negative effect on saving and a positive effect on borrowing. However, the saving can be affected in two other ways. Because income uncertainty reduced by insurance, people have less precautionary incentive to save, in this sense, saving tends to decrease. At the same time, if we assume that people have rational expectations, as they are expected to become richer in future periods, they will smooth consumption across periods by increasing consumption and reducing saving in the current period. Furthermore, if the purchase of insurance is subsidized, it has a positive effect on farmers' wealth, which has a positive effect on saving.

Consider a representative farmer who lives for two periods with initial wealth W_0. In the first period, the farmer consumes C_1 and uses the remaining wealth for investment. There are two ways to invest this money: one is to save it in the bank with a saving interest rate R_f, the other is to invest it in a risky project like crop production which has a return function $F(\cdot)$. The farmer can borrow from a local bank for investment in a risky project with interest rate R_B. So the total investment I on the risky project includes the initial wealth less consumption and saving, and a loan equal to B from the bank. The return of the risky project is uncertain because it depends on whether a disaster happens in period one. In this simple model I assume that there are two states: a good state (no disaster) and a bad state (disaster). In the good state, the farmer gets $F(I)$, while in bad state he gets

[5]For example, consider a farmer who has 5 mu in tobacco production. If the normal yield per mu is 500kg and because of a windstorm, the farmer's yield decreased to 250kg per mu, then the loss rate is 50% and he will receive 420*50% = 210 RMB per mu from the insurance company.

[6]Throughout the model I assume that farmers who are provided with insurance buy it compulsory in every period, while those who are not provided with insurance cannot buy it in any period.

nothing. Assume that there is no strategic default and that farmers have limited liability, then in the good state, the farmer will repay fully in the second period; under a bad state, the farmer default on the loan if he does not have money to repay.

Suppose that for a farmer who invests I on the risky project (production), in order to buy an insurance which covers all his production[7], he needs to pay a premium which equals δI[8]. The production insurance works as follows: in the bad state, the farmer will be reimbursed by the insurance company by an amount equals to part of the cost invested in the risky project, γI. As a result, even in the bad state, the farmer who purchased insurance will be able to repay part or all of the loan.

In order to compare farmers' financial and investment behavior depending on whether they have insurance or not, I will solve the two-period model separately for insured and uninsured farmers because in the second period, their consumptions are different in the bad state. Throughout the model I assume that farmers are price takers: they don't think their behavior can influence either the premium charged by the insurance company or the saving and borrowing interest rate set by the bank.

3.3.1 Two-period model when insurance is not provided

The optimization problem as follows:

$$\begin{gathered} \max_{C_1,I,B} U(C_1) + E\beta U(C_2) \\ \iff \max_{C_1,I,B} U(C_1) + \beta p U\left[F(I) - (1+R_B)B + (1+R_f)S\right] + \beta(1-p)U\left[(1+R_f)S\right] \\ \text{s.t. } I = W_0 - C_1 - S + B \end{gathered}$$

In order to solve the model, I need to make assumptions about the functional forms of the return function and the utility function. Here I assume that the return function and the utility function are:

$$\begin{gathered} F(I) = I^\alpha, \alpha < 1^9 \\ U(C) = \log C \end{gathered}$$

Then the first order conditions are:

$$U'(C_1) = \beta p U'\left[F(I) - (1+R_B)B + (1+R_f)S\right] F'(I) = \beta p U'(C_g) F'(I) \tag{3.1}$$

$$\beta p U'(C_g)\left[(1+R_f) - F'(I)\right] + \beta(1-p)U'\left[(1+R_f)S\right](1+R_f) = 0 \tag{3.2}$$

$$\beta p U'(C_g)\left[F'(I) - (1+R_B)\right] = 0 \tag{3.3}$$

[7] An assumption here is that to reduce the average risk and to prevent adverse selection, the insurance company requires the farmer to buy insurance for all his production area.

[8] In my data, δ should be quite low because farmers only need to pay 6 RMB per mu to buy the insurance, but the production cost (I) is around 400-600 RMB per mu.

[9] This return function form can exclude the case of infinite investment.

$$\Rightarrow F'(I^*) = 1 + R_B \quad (3.4)$$

According to the return function form, I can rewrite equation (3.4) as:

$$F'(I^*) = \alpha I^{*\alpha-1} = 1 + R_B$$
$$\Rightarrow I^* = \left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}} \quad (3.5)$$

So the optimal level of investment is decreasing in the borrowing interest rate R_B, or in other words, people tend to investment more on the risky project when the cost of borrowing is lower. Part 1 in Appendix C gives the solution of the above optimization problem.

3.3.2 Two-period model when insurance is provided

If the farmer has production insurance, the framework is as follows:

$$\max_{C_1,B,S} U(C_1) + \beta pU[C_g] + \beta(1-p)U[C_b]$$
$$s.t. I = B + [W_0 - C_1 - S - \delta I]$$
$$\Rightarrow I = \frac{W_0 - C_1 - S}{1+\delta} + \frac{B}{1+\delta}$$

Where C_g and C_b are the farmer's consumption in period two under good and bad state, respectively. The biggest difference in this model is that under bad state, the farmer receives a reimbursement from the insurance company which covers part of their cost, which equals $\gamma I = \gamma\frac{W_0 - C_1 - S}{1+\delta} + \gamma\frac{B}{1+\delta}$, so I can write the return of production under bad state as γI. Since I have assumed there's no strategical default, the farmer will repay the bank $\gamma\frac{B}{1+\delta}$, which is the return that is generated by a loan with size B. Given this, the consumption in period two under two states is defined as follows, respectively:

$$C_g = F(I) - (1+R_B)B + (1+R_f)S$$
$$C_b = \frac{\gamma}{1+\delta}(W_0 - C_1 - S + B) - \frac{\gamma}{1+\delta}B + (1+R_f)S$$

The three first order conditions are:

$$U'(C_1) - \beta pU'(C_g)F'(I)\frac{1}{1+\delta} - \beta(1-p)U'(C_b)\frac{\gamma}{1+\delta} = 0 \quad (3.6)$$
$$\beta pU'(C_g)\left[-(1+R_B) + F'(I)\frac{1}{1+\delta}\right] = 0 \quad (3.7)$$
$$\beta pU'(C_g)\left[(1+R_f) - F'(I)\frac{1}{1+\delta}\right] + \beta(1-p)U'(C_b)\left[-\frac{\gamma}{1+\delta} + 1 + R_f\right] = 0 \quad (3.8)$$

The utility and return function forms are the same as that in previous sections:

$$U(C) = \log C$$
$$F(I) = I^\alpha, 0 < \alpha < 1$$

Part 2 in Appendix C gives the solution of the above optimization problem.

[10]This makes sense since project has return only in good states and it is the only time repayment is required.

3.3.3 Combine the two models

The expressions of the optimal investment, consumption, saving and borrowing for insured and uninsured farmers as follows:

$$I^*(insured) = \left(\frac{(1+R_B)(1+\delta)}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

$$I^*(unisured) = \left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

$$C_1^*(insured) = \frac{1}{D+E}\left[\frac{(R_B-R_f)\gamma+(1+R_B)\left[(1+\delta)(1+R_f)-\gamma\right]}{(1+R_B)(1+\delta)\left[(1+\delta)(1+R_f)-\gamma\right]}W_0 + (\alpha^{-1}-1)(\frac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}}\right]$$

$$C_1^*(uninsured) = \frac{1}{1+\beta}\left[W_0 + (\alpha^{-1}-1)\left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]$$

$$S^*(insured) = \frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}\frac{(R_B-R_f)\gamma+(1+R_B)\left[(1+\delta)(1+R_f)-\gamma\right]}{(1+R_B)(1+\delta)\left[(1+\delta)(1+R_f)-\gamma\right]}W_0$$

$$+\frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}(\alpha^{-1}-1)(\frac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}} - \frac{\gamma W_0}{(1+R_f)(1+\delta)-\gamma}$$

$$S^*(unisured) = \frac{(1+R_B)(1-p)\beta}{(1+\beta)(R_B-R_f)}\left[W_0 + (\alpha^{-1}-1)\left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]$$

$$B^* = (1+R_B)^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{\alpha}{\alpha-1}}\alpha^{-\frac{\alpha}{\alpha-1}} - \frac{D}{1+R_B}C_1^* + \frac{1+R_f}{1+R_B}S^*$$

$$B^*(unisured) = (1+R_B)^{\frac{1}{\alpha-1}}\alpha^{-\frac{\alpha}{\alpha-1}} - \frac{\beta\left[p(R_B+1)-(1+R_f)\right]}{R_B-R_f}C_1^*$$

3.3.4 Break-even conditions of the bank

Now I have solved farmers' optimization problem, the next step is to consider the break-even conditions of the bank[11].

If the bank's client does not have insurance, he gets nothing in bad state, so the break-even condition is:

$$B(1+R_f) = p(1+R_B)B$$
$$\Rightarrow R_B = [1+R_f]\frac{1}{p} - 1$$

If insurance is purchased, the break-even condition becomes:

$$(1+R_f)B = p(1+R_B)B + (1-p)\frac{\gamma}{1+\delta}B$$
$$\Rightarrow R_B = \left[1+R_f-\frac{(1-p)\gamma}{1+\delta}\right]\frac{1}{p} - 1.$$

In summary:

$$R_B = \begin{cases} [1+R_f]\frac{1}{p} - 1, & \text{if not insured} \\ \left[1+R_f-\frac{(1-p)\gamma}{1+\delta}\right]\frac{1}{p} - 1, & \text{if insured} \end{cases}$$

We can see that the bank will set a lower interest rate for people who have insurance because their repayments are better guaranteed.

[11] Here I assume that the institution's objective is to break-even for simplicity.

3.3.5 Conclusion of the model

Now I plug the interest rate into optimal decisions in 3.3.3 and compare the magnitude of investment, consumption, saving and borrowing between insured and uninsured farmers.

- Investment: Farmers will invest more when they have insurance

$$I^*(insured) = \left(\frac{(1+R_B)(1+\delta)}{\alpha}\right)^{\frac{1}{\alpha-1}} = \left(\frac{\frac{1+R_f}{p} - \frac{(1-p)\gamma}{(1+\delta)p}}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

$$I^*(unisured) = \left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}} = \left(\frac{\frac{1+R_f}{p}}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

Because $\alpha - 1 < 0$, so if $\frac{(1-p)\gamma}{(1+\delta)p} > 0$, the investment increase as a result of insurance provision. Intuitively, when insurance is provided, borrowing will become cheaper and the expected return of the risky project will increase, so investing in the risky project becomes more attractive.

- Consumption: The first period consumption is higher when the farmer have insurance.

$$C_1^*(insured) = C_1^*$$
$$= \frac{1}{D+E}\left[\frac{(R_B-R_f)\gamma+(1+R_B)\left[(1+\delta)(1+R_f)-\gamma\right]}{(1+R_B)(1+\delta)\left[(1+\delta)(1+R_f)-\gamma\right]}W_0 + (\alpha^{-1}-1)\left(\frac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}}\right]$$

$$C_1^*(unisured) = \frac{1}{1+\beta}\left[W_0 + (\alpha^{-1}-1)\left(\frac{R_f/p+1/p}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]$$

Because $\frac{1+R_B}{1+R_f} + \frac{(R_B-R_f)\gamma}{(1+R_f)[(1+R_B)(1+\delta)-\gamma]} > 1$, $\left(\frac{R_f/p+1/p-(1-p)\gamma/p(1+\delta)}{\alpha}\right)^{\frac{1}{\alpha-1}} > \left(\frac{R_f/p+1/p}{\alpha}\right)^{\frac{1}{\alpha-1}}$
and $(1+R_f)(1-p)(1+\delta-\delta\eta) > R_f\delta\eta$[12]
then $C_1^*(insured) > C_1^*(unisured)$

So the second message from the model is that, people who bought insurance will consume more in the first period. This makes sense because if a farmer has insurance, he expect himself to be richer in the second period compared to the condition when he does not have insurance, so he will smooth the consumption between periods by increasing the consumption in period one.

- Saving: The provision of insurance can decrease farmers' total saving and saving rate in period one.

[12]This condition holds in my data.

$$S^*(insured) = \frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}\frac{(R_B-R_f)\gamma+(1+R_B)\left[(1+\delta)(1+R_f)-\gamma\right]}{(1+R_B)(1+\delta)\left[(1+\delta)(1+R_f)-\gamma\right]}W_0 +$$
$$\frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}(\alpha^{-1}-1)(\frac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}} - \frac{\gamma W_0}{(1+R_f)(1+\delta)-\gamma}$$
$$= \left[\frac{\beta}{1+\beta} - \frac{\beta\gamma p}{(1+\beta)\left[(1+\delta)(1+R_f)-\gamma\right]}\right]W_0 + (\alpha^{-1}-1)(\frac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}}\frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}$$
$$S^*(unisured) = \frac{\beta}{(1+\beta)}\left[W_0 + (\alpha^{-1}-1)\left(\frac{1/p+R_f/p}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]$$
$$= \frac{\beta}{1+\beta}W_0 + (\alpha^{-1}-1)\alpha^{-\frac{1}{\alpha-1}}(\frac{1}{p}+\frac{R_f}{p})^{\frac{1}{\alpha-1}}\frac{\beta}{1+\beta}$$

Because $\left[\frac{\beta}{1+\beta} - \frac{\beta\gamma p}{(1+\beta)\left[(1+\delta)(1+R_f)-\gamma\right]}\right] < \frac{\beta}{(1+\beta)}$, so if W_0 is large enough, $S^*(insured) < S^*(unisured)$ and $Savingrate^*(insured) < Savingrate^*(unisured)$. This result is consistent with the precautionary saving story: farmers' future income uncertainty is decreased by introducing insurance, so people have less precautionary incentive to save in the first period for smoothing future consumption.

- Borrowing: The effect of insurance provision on borrowing is ambiguous.

The total investment on risky project is $I = B + [W_0 - C_1 - S]$, I have proved that the provision of insurance will increase C_1 and I, and decrease S, so the effect on B is ambiguous.

In summary, the conclusion from this two-period model is that insurance has a positive effect on investment in risky projects and consumption, and it reduces farmers' total saving and saving rate. As a result, its effect on borrowing is not determined.

3.4 Data and Summary Statistics

As shown in Table 3.1, the empirical analysis is based on data from 8 tobacco production counties in Jiangxi province of China: Guangchang, Shicheng, Ganxian, Huichang, Xinfeng, Xinguo, Ruijin, and Quannan. Among these 8 counties, only tobacco farmers in Guangchang were eligible to buy the tobacco insurance policy after 2002. In Guangchang county, only tobacco households whose main source of income is from tobacco production were offered with production insurance, while households working on other activities were not eligible to buy similar products.

The primary data source is the household level panel dataset, ranging from 2000 to 2008, provided by the Rural Credit Cooperatives (RCC). The whole sample includes around 6500 households. The data is composed of two parts. The first part is the administrative of RCC, including saving and borrowing information of their clients[13]. Specifically, it has variables such as loan certification number, total borrowing during the year, interest rate, use of loan,

[13]since more than 90% farmers are RCC clients in Jiangxi province, this data is representative of the whole sample of farmers.

repayment, total annual saving, savings in the deposit account, savings in the current account, and annual growth in savings[14]. The second part is RCC annual survey data[15], which contains family background information: age, gender, occupation and education of household heads, primary and secondary source of household income, family address, household size, and labor size; household income and production: including total annual income, household income from different sources, remittance, area of land for cultivation, and production areas of different crops.

The data includes 6548 households in total, among which 3580 households are tobacco households, and 2968 households are other households whose main source of income is not tobacco production[16]. For tobacco households, 1429 of them are in the treatment region where insurance policy was available, and 2151 of them are in control regions.

The summary statistics of key variables before the insurance policy was implemented (2000-2002) are provided in Table 3.2. Household heads are almost exclusively male, and the average age is around 40. The average household size includes around 5 people, and household heads have an average education level of between primary and secondary schools. The above household characteristics are very similar across different household groups. The average annual household income of tobacco households in treatment regions equals 10,650 RMB, while that of tobacco households in control regions is a bit higher, of around 12,000 RMB. Annual income of non-tobacco households is much lower, with only 7,270 RMB. Consider households' borrowing behavior, the average borrowing of non-tobacco households is the highest (4,980 RMB), followed by tobacco households in control regions (4,560 RMB), and tobacco households in treatment regions (3,900 RMB). The household saving rate is defined as the ratio between net annual saving and household income. For tobacco households in treatment regions, their saving rate is around 3.6%, and it is lower than that of tobacco households in control regions (4.5%). Saving rate of non-tobacco households is similar as that of tobacco households in treatment regions, of around 3.4%[17]. This table suggests that as treatment and control tobacco households behave statistically different in pre-policy periods, I cannot study the policy impact by taking simple difference.

In order to check whether Difference-in-Difference (DD) estimation can be a convincing strategy in this context, I test the common-trend assumption in Table 3.3, using the following regression:

[14]While RCC is the main place for farmers to make deposit, households may have saving accounts in other institutions. As a result, the amount of saving in RCC does not represent a household's total saving. To account for this factor, RCC reported village-level ratio of RCC saving in total household saving, and I adjusted the RCC saving data by the ratio in all of the empirical analysis.

[15]RCC implements household survey every year in order to adjust the lending interest rate and loan ceiling for each household.

[16]These households work on agricultural activities such as rice production, cultivation, etc. or non-agricultural activities.

[17]Households with outliers (the lowest or higest 1%) in income, loan size, and savings were deleted from the sample for analysis.

$$Y_{irt} = \eta_0 + \eta_1 Year_t + \eta_2 Insurance_{ir} + \eta_3 Year_t * Insurance_{ir} + \epsilon_{irt} \quad (3.1)$$

Where i, r, t are household, region, and year indices respectively. $Insurance_{ir}$ is the treatment indicator equal to 1 for treatment regions (guangchang county) and 0 for control regions (other counties). The common-trend assumption does not hold if coefficient of the interaction term, η_3, is statistically significant. Results show that the common trend assumption is not valid for all outcomes I'm interested in, so using only DD estimation is not sufficient.

To get a basic sense about how insurance provision impacts production, borrowing, and saving, I plot the evolution of these variables in Figures 3.1 to 3.5. Figure 3.1 shows that while tobacco production was similar between tobacco households in treatment and control regions, the production scale increased a lot in treatment regions after 2002. Refer to Figure 3.2, we can see that while tobacco households in treatment regions borrowed less than those in control regions before 2002, the pattern reversed after 2003. However, Figure 3.3 shows that borrowing pattern is different across the sample period between non-tobacco households in treatment and control regions, which suggests that there might be some regional-specific trend we should control for when estimating the policy effect. In Figure 3.4, I show that for tobacco households, while saving rate is higher in control regions compared with that in treatment regions, the trend reversed a little bit after 2004. The difference in saving rates of non-tobacco households between treatment and control regions is much larger, as described in Figure 3.5.

Table 3.4 reports the average area of tobacco production, size of loans, and saving rate by time period, region, and sector eligibility. Take loan size for example. For each region-sector category, the average loan size increases from the period of 2000-2002 to the period 2003-2008, reflecting the aggregate economic trend. For tobacco households, the average loan size in treatment regions increases by 1,450 RMB more than that of households in control regions. This can be a result of both the implementation of insurance policy and other region-specific changes. Actually for non-tobacco households, the size of loan in treatment regions also grows faster than households in control regions by 1,080 RMB. Taking into account this regional difference in the absence of the insurance policy, the loan size of tobacco households in treatment regions increases by 970 RMB more than that of similar households in control regions. The regression analysis in the next section demonstrates that this effect is robust to controlling for other confounding factors. These results suggest that triple difference estimation can be a more convincing empirical strategy compared with DD in this case.

3.5 Estimation Strategies and Results

3.5.1 Empirical Strategies

The implementation of tobacco insurance policy introduced variations in insurance provision on three dimensions: years before and after the policy was introduce, regions with and without the policy, and eligible and ineligible households (tobacco households v.s. other households). These variations allow me to use both difference-in-difference (DD) and difference-in-difference-in-difference (DDD) estimation as the empirical strategy. First, the DD analysis compares the change in tobacco household's behavior in treatment regions before and after 2002 with that of tobacco households in control regions, assuming that tobacco households in treatment and control regions follow the same trend in the absence of the provision of insurance policy. The estimation equation is as follows:

$$Y_{irt} = \alpha_0 + \alpha_1 After_{it} + \alpha_2 Insurance_{ir} + \alpha_3 After_{it} * Insurance_{ir} + \epsilon_{irt} \tag{3.2}$$

Where i, r, t are household, region, and year indices respectively. This framework is based on tobacco households only. Y represents outcome variables including tobacco production area, size of loan borrowed from the rural bank, and saving rate. $After$ is a dummy variable equal to 1 for the 2000-2002 period and 0 for years 2003-2008, which reflects the impact of time-varying aggregate economic environment and policies on outcomes. $Insurance_{ir}$ is the treatment indicator equal to 1 for treatment regions (guangchang county) and 0 for control regions (other counties). The coefficient of interest is the one before the interaction term between $After$ and $Insurance_{ir}$, α_3.

However, the DD estimation cannot remove all confounding factors. For example, there may be some other contemporary changes in economic environment or policies specific to the treatment region that can influence households' production and financial decisions. This can be captured by taking another DD analysis, which compares behavior of non-tobacco households in treatment regions before and after 2002 with that of non-tobacco households in control regions. As a result, the DDD framework, which takes the difference between the two differences from the first two steps, can further control for region-specific trend. Under the DDD framework, we don't need to assume that behaviors of tobacco households in both treatment and control regions evolve similarly in expectation, but only need to assume that the difference affect tobacco households and other households similarly (in other words, there are no other region-sector specific policy changes). I will test this assumption later. The DDD regression is as follows:

$$\begin{aligned} Y_{ijrt} = \beta_0 &+ \beta_1 After_{it} + \beta_2 Insurance_{ir} + \beta_3 Tobacco_{ij} + \beta_4 After_{it} * Insurance_{ir} \\ &+ \beta_5 After_{it} * Tobacco_{ij} + \beta_6 Tobacco_{ij} * Insurance_{ir} \\ &+ \beta_7 After_{it} * Insurance_{ir} * Tobacco_{ij} + \epsilon_{ijrt} \end{aligned} \tag{3.3}$$

Where j is sector indicator, and $Tobacco_{ij}$ is a dummy variable equal to 1 for tobacco households and 0 otherwise. The coefficient of the time, region, and sector interaction (β_7) captures the average effect of insurance provision on household behavior, after other confounding factors are removed.

Significant influences of insurance provision on households' production and investment decisions may take place either shortly after the policy was introduced or several years after that, and the magnitude of effect may change as time evolve. Consequently, it would be interesting to test the dynamic effect of insurance provision on household behavior. The estimation equation is as follows:

$$\begin{aligned} Y_{ijrt} = \rho_0 &+ \rho_1 Year_t + \rho_2 Insurance_{ir} + \rho_3 Tobacco_{ij} + \rho_4 Year_t * Insurance_{ir} \\ &+ \rho_5 Year_t * Tobacco_{ij} + \rho_6 Tobacco_{ij} * Insurance_{ir} \\ &+ \rho_7 Year_t * Insurance_{ir} * Tobacco_{ij} + \epsilon_{ijrt} \end{aligned} \tag{3.4}$$

Where $Year_t$ includes a set of year dummies. Estimating the above equation can not only allow me to test the dynamic effect, but also to test the crucial assumption which validates the DDD estimation: in the absence of the insurance policy, the production and financial behaviors of tobacco households and non-tobacco households should evolve similarly.

The magnitude of impact of insurance provision on household behavior can be different for different groups of households. I consider two types of heterogeneity here, depending on farming size and the importance of migration remittance in household income. The regression is as follows:

$$\begin{aligned} Y_{ijrt} = \gamma_0 &+ \gamma_1 After_{it} + \gamma_2 Insurance_{ir} + \gamma_3 Tobacco_{ij} + \gamma_4 After_{it} * Insurance_{ir} \\ &+ \gamma_5 After_{it} * Tobacco_{ij} + \gamma_6 Tobacco_{ij} * Insurance_{ir} \\ &+ \gamma_7 After_{it} * Insurance_{ir} * Tobacco_{ij} + \gamma_8 Index_{it} + \gamma_9 Index_{it} * After_{it} \\ &+ \gamma_{10} Index_{it} * Insurance_{ir} + \gamma_{11} Index_{it} * Tobacco_{ij} \\ &+ \gamma_{12} Index_{it} * After_{it} * Insurance_{ir} + \gamma_{13} Index_{it} * After_{it} * Tobacco_{ij} \\ &+ \gamma_{14} Index_{it} * Insurance_{ir} * Tobacco_{ij} \\ &+ \gamma_{15} Index_{it} * After_{it} * Insurance_{ir} * Tobacco_{ij} + \epsilon_{ijrt} \end{aligned} \tag{3.5}$$

Where $Index_{it}$ is an indicator equal to 1 if in the pre-policy period (2000-2002), the households' total production area or %migration remittance in total income is higher than the sample median, and 0 otherwise. The coefficient of interest is γ_{15}.

3.5.2 Estimation Results

Tables 3.5 - 3.8 report DD and DDD estimation results on the effect of insurance provision on households' production, borrowing, and saving decisions, respectively[18]. Look at the effect on production first. Refer to Column (1) in Table 3.5, the increase in tobacco production post of 2002 is 1.161 mu larger for households in treatment regions compared with households in control regions. Since the pre-policy mean of tobacco production in treatment regions is around 5.25 mu (refer to Table 3.2), it means insurance provision can raise tobacco production by 22%. This is consistent with the story that as the expected return of tobacco production increases once insurance was provided, it gives household more incentive to invest more on tobacco production. Column (2) includes year dummies in addition, and the magnitude of effect increased a little bit. In Column (3), I further control household characteristics, including household size, annual household income, age and education of household head. The magnitude of treatment effect remains similar, of around 1.2 mu (23%). Column (3) also shows that household with higher annual income tend to produce tobacco more, as the production cost of tobacco cultivation is high relative to that of other crops. Moreover, larger households, and those with more well-educated and younger household heads, are likely to have a larger production scale. This can be explained by the fact that tobacco production not only requires more labor than other productions, but also good knowledge and techniques to have high yield and good quality tobacco.

In Table 3.6, I look at the impact of insurance provision on households' choice of production diversification, which is defined as one minus the Herfindahl index of agricultural production. The results show that agricultural production became less diversified after the insurance was provided, by around 29%. This means that households tend to focus more on producing the insured crop after the intervention.

Second, Table 3.7 reports the DDD estimation results on the effect of insurance policy on borrowing. Results suggest a significant insurance treatment effect on borrowing of around 972 RMB. Comparing it to the average loan size of tobacco households in treatment regions before 2003 (shown in Table 3.2), it tells us that tobacco households borrow 25% more once their production is insured.

Third, the effect of insurance provision on household saving is reported in Table 3.8. According to Columns (1) and (2), after tobacco insurance policy was introduced, the increase in the average saving rate of tobacco households in treatment regions is around 1.24 percentage points lower than that of tobacco households in control regions. This means that providing insurance can decrease household saving rate by more than 30%. In Columns (3) and (4), I consider the level of net saving rather than the saving rate. The estimation results show that while the insurance policy has a significant impact on saving rate, it did not influence the level of saving significantly. Finally, in Columns (5) and (6), I estimate the effect of insurance on the composition of saving. In China, households can have two types

[18]Please note that the DDD framework is not applicable to estimating the effect on tobacco production area, because there's almost no tobacco production for non-tobacco households

of saving accounts: fixed-term saving or flexible-term saving (like the checking accounts in the US). I show that the insurance policy does not have any statistically significant impact on the composition of saving.

The dynamic impact of insurance provision on households' borrowing and saving behavior is illustrated in Table 3.9. The result shows that first, before the insurance policy was introduced, there's no significant difference between households with or without tobacco production, because interactions of 2001-2002 year dummies, region, and sector are insignificant. Second, according to Column (1), the effect of insurance provision on borrowing is insignificant until 2 years after the intervention. But both the magnitude and significance of the effect persists through the end of our sample period. In contrast, according to Column (2), insurance impact on household saving became significant three after the policy was introduced, but the magnitude and significance decreases and became insignificant towards the end of the sample period.

In Table 3.10, I report the heterogeneity in the impact of insurance, depending on how large the farming size is, and how important is the migration remittance for the household income. Columns (1) - (3) shows that insurance provision has a larger effect on borrowing for large farmers, while the effect on production and saving is not statistically different for farmers with different farming size. In Columns (4) - (6), I show that the effect of insurance policy has a smaller impact on the production and borrowing decisions of households who depend more on migration remittance.

Once the insurance policy was implemented on tobacco farmers, we may expect endogenous switch of non-tobacco households to tobacco households. If a significant number of households do so, the effect might be overestimated. In Table 3.11, I report the percentage of households that stay in the same sector, switch from tobacco to non-tobacco sector, and switch from non-tobacco sector to tobacco sector between the previous and the current year for treatment and control regions. It shows that only a very small fraction of households changed sectors during the sample period. I did a robustness check by excluding all households who had ever switched sectors and it does not change the effect much.

3.6 Conclusions

Household incomes in developing rural economies are subject to great uncertainty. As a result, many developing countries are making efforts to improve the quality and coverage of agricultural insurance products. Taking advantage of a natural experiment of insurance provision in rural China, this paper uses both DD and DDD estimations to study the effect of insurance provision on households' production and financial decisions. I find that households tend to extend tobacco production size once it is insured. Moreover, insurance provision not only make households borrow more from bank, but also decrease household saving rate. However, while insurance impact on borrowing persists in the long-run, that on saving is only significant in the medium-run and vanishes in the long-run.

Figure 3.1. Evolution of Tobacco Production, by Treatment

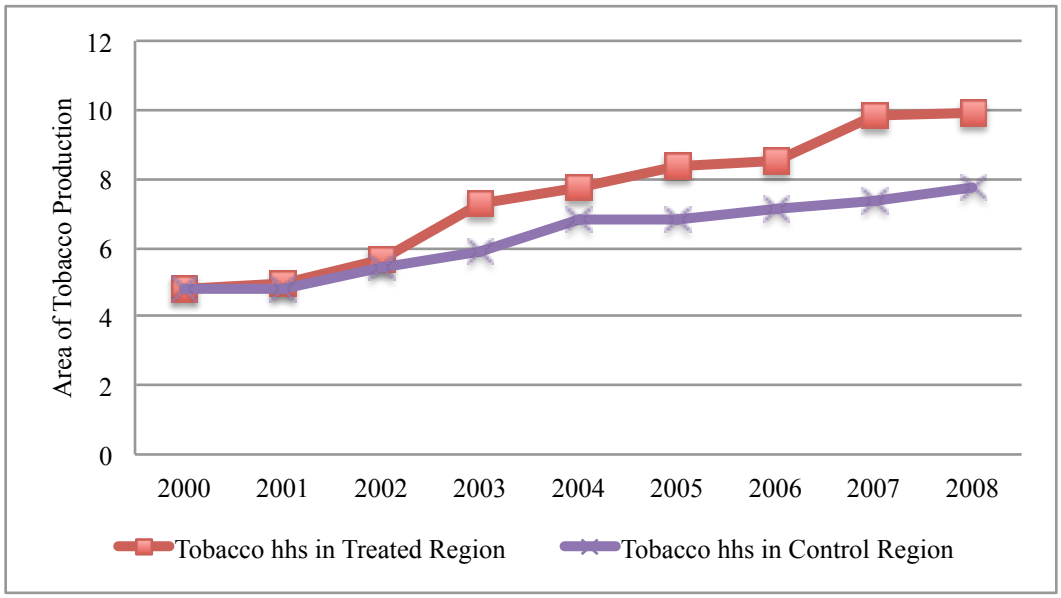


Figure 3.2. Evolution of Loan Size for Tobacco Households, by Treatment

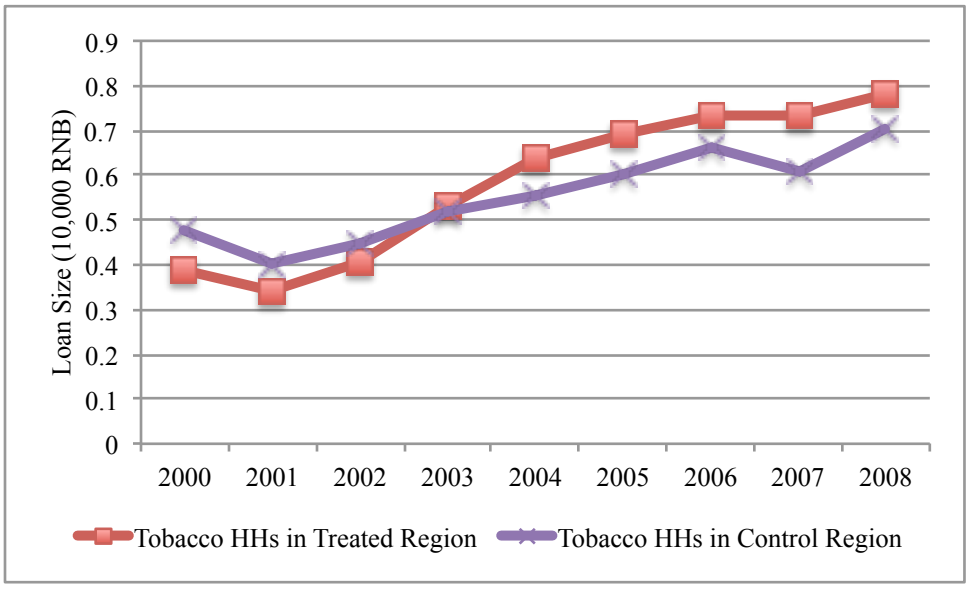


Figure 3.3. Evolution of Loan Size for Other Households, by Treatment

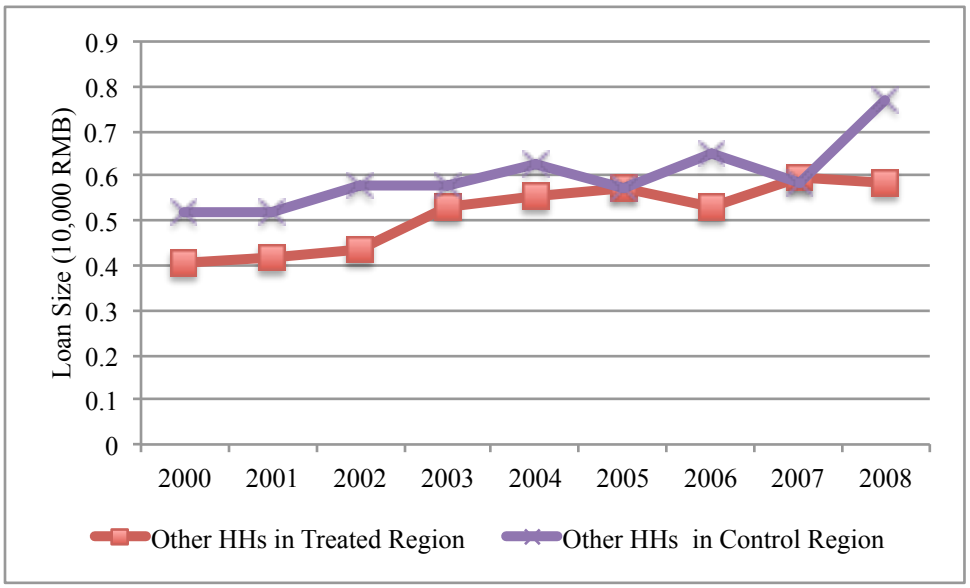


Figure 3.4. Evolution of Saving for Tobacco Households, by Treatment

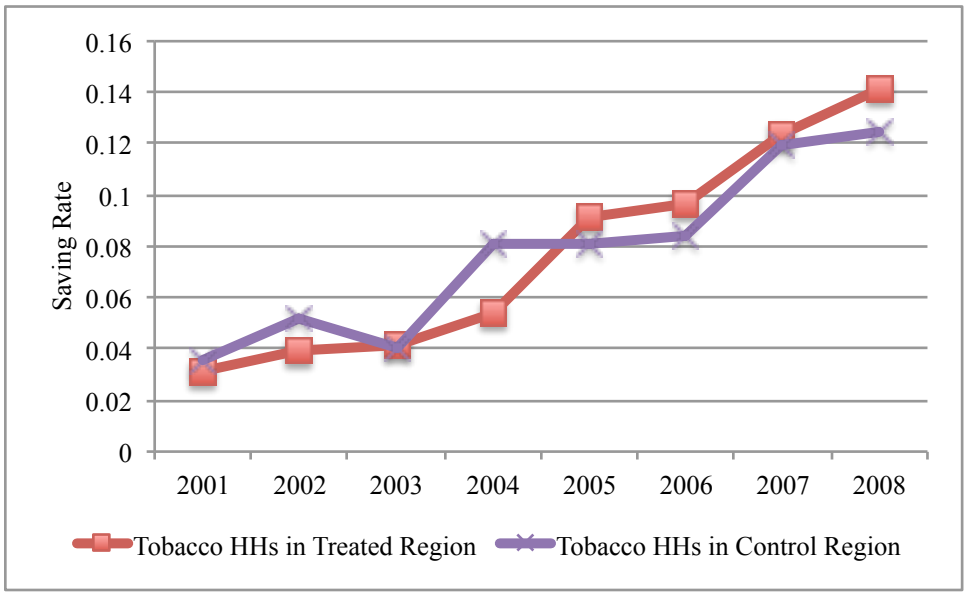


Figure 3.5. Evolution of Saving for Other Households, by Treatment

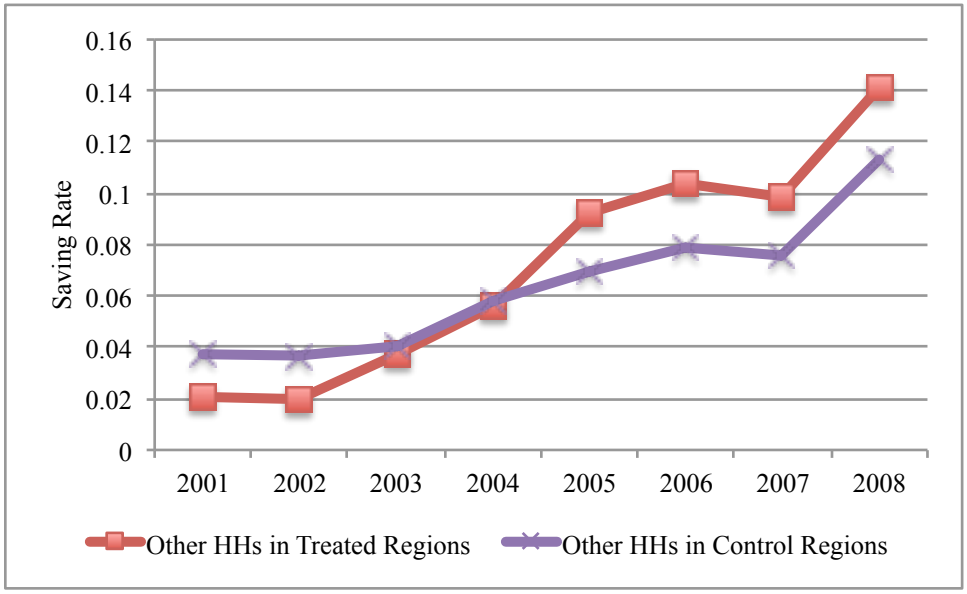


Table 3.1. Insurance Provision in Tobacco Production Counties of Jiangxi Province

County	Insurance Provision	Start Year	Premium	Subsidy of Premium	Maximum Payout
Guangchang	Yes	2003	12	50%	420
Shicheng	No				
Ningdu	No				
Ganxian	No				
Huichang	No				
Xinfeng	No				
Xingguo	No				
Ruijin	No				

Notes: The unit of premium and payouts is RMB per mu (1 mu = 1/15 hectare). The exchange rate between US dollars and RMB is around 6.3.

Table 3.2. Summary Statistics

	Tobacco Households			Other Households	All Sample
	Treated	Control	Difference		
	(1)	(2)	(3)	(4)	(5)
Number of Households	*1429*	*2151*		*2968*	*6548*
Gender of Household Head (1 = Male, 0 = Female)	0.996	0.982	0.014***	0. 978	0.983
	(0.062)	(0.134)	(0.000)	(0.146)	(0.131)
Age	40.418	40.731	-0.313*	40.205	40.429
	(8.959)	(8.124)	(0.091)	(8.645)	(8.526)
Household Size	4.781	4.728	0.053*	4.930	4.832
	(1.022)	(1.355)	(0.054)	(1.312)	(1.284)
Education (0 = illiteracy, 1 = primary,	1.626	1.759	-0.133***	1.813	1.760
2 = secondary, 3 = high school, 4 = college)	(0.54)	(0.929)	(0.000)	(0.644)	(0.746)
Area of Tobacco Production (mu)	5.249	4.999	0.249***	0.307	2.857
	(2.119)	(2.874)	(0.000)	(1.194)	(3.175)
Production Diversification Index (0-1)	0.389	0.275	0.114***	0.119	0.237
	(0.229)	(0.261)	(0.000)	(0.203)	-0.256
Annual Household Income (10,000 RMB)	1.065	1.202	-0.137***	0.727	0.956
	(0.477)	(1.402)	(0.000)	(0.941)	(1.094)
Loan Size (10,000 RMB)	0.390	0.456	-0.066***	0.498	0.483
	(0.203)	(0.189)	(0.003)	(0.089)	(0.13)
Saving Rate (Net Saving Divided by Income)	0.036	0.045	-0.009***	0.034	0.038
	(0.079)	(0.12)	(0.003)	(0.093)	(0.101)

Notes: This table reports the mean of key variables in pre-treatment periods (2000-2002). For columns (1), (2), (4) and (5), standard deviations are in brackets. For column (3), P-value for F test of equal means of two groups are in brackets.

Table 3.3. Test Common Trend in Key Outcome Variables Before Policy Intervention

VARIABLES	Area of Tobacco Production (mu)	Loan Size (10,000 RMB)	Saving Rate (Net Saving Divided by Income)
	(1)	(2)	(3)
Year	0.322	-0.0543**	0.0163**
	(0.305)	(0.0233)	(0.00776)
Insurance	0.123	-0.110***	0.00429
(= 0 if control region, = 1 if treatment region)	(0.305)	(0.0161)	(0.0383)
Year*Insurance	0.160	0.0708***	-0.00860
	(0.308)	(0.0233)	(0.00776)
Observations	9,201	659	5,761
R-squared	0.080	0.034	0.006

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4. Area of Tobacco Production, Loan Size, and Saving Rate by Region, Sector, and Year

	Tobacco Households			Other Households		
	2000-2002	2003-2008	Difference	2000-2002	2003-2008	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
I. Area of Tobacco Production (mu)						
Treatment	5.249	8.464	3.215***			
	(0.038)	(0.043)	(0.000)			
Control	4.999	7.054	2.054***			
	(0.037)	(0.027)	(0.000)			
DD			1.161***			
			(0.000)			
II. Loan Size (10,000 RMB)						
Treatment	0.390	0.724	0.004***	0.412	0.568	0.156***
	(0.022)	(0.016)	(0.000)	(0.006)	(0.011)	(0.000)
Control	0.456	0.645	0.189***	0.523	0.630	0.108***
	(0.008)	(0.013)	(0.000)	(0.002)	(0.032)	(0.000)
DD			0.145***			0.048
			(0.000)			(0.163)
DDD			0.097**			
			(0.036)			
III. Saving Rate (Net Saving Divided by Income)						
Treatment	0.036	0.086	0.049***	0.020	0.098	0.078***
	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.000)
Control	0.045	0.093	0.049***	0.037	0.064	0.028***
	(0.002)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)
DD			0.001			0.05***
			(0.804)			(0.000)
DDD			-0.05***			
			(0.000)			

Notes: For columns (1), (2), (4) and (5), standard deviations are in brackets. For columns (3) and (6), P-value are in brackets.

Table 3.5. Effect of Insurance Provision on Production

VARIABLES	Area of Tobacco Production (mu)		
	(1)	(2)	(3)
After	2.054***	7.938***	7.279***
(= 0 if 2000-2002, = 1 if 2003-2008)	(0.320)	(2.261)	(1.938)
Insurance	0.249	0.175	0.338***
(= 0 if control region, = 1 if treatment region)	(0.370)	(0.195)	(0.112)
After * Insurance	1.161***	1.450***	1.223***
	(0.320)	(0.167)	(0.116)
Household Size			0.0727***
			(0.00790)
Annual Household Income (10,000 RMB)			0.787***
			(0.257)
Education (0 = illiteracy, 1 = primary,			0.343*
2 = secondary, 3 = high school, 4 = college)			(0.178)
Age			-0.0252**
			(0.0112)
No. of Observation	31,207	31,207	31,207
Year Fixed Effects	No	Yes	Yes
R-squared	0.105	0.131	0.226

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.6. Effect of Insurance Provision on Production Diversification

VARIABLES	Production Diversification (0-1)		
	(1)	(2)	(3)
After	0.0543	0.0364	0.0268
(= 0 if 2000-2002, = 1 if 2003-2008)	(0.0536)	(0.063)	(0.0550)
Insurance	-0.0492	-0.0525	-0.0470
(= 0 if control region, = 1 if treatment region)	(0.0753)	(0.0524)	(0.0759)
Tobacco Household	0.144**	0.144***	0.145***
(= 0 if No, = 1 if Yes)	(0.0585)	(0.0214)	(0.0491)
After * Insurance	0.0755	0.0670	0.0657
	(0.0536)	(0.0607)	(0.0612)
After * Tobacco Household	-0.0320	-0.0433	-0.0482
	(0.0320)	(0.0386)	(0.0336)
Tobacco Household * Insurance	0.164***	0.168***	0.171***
	(0.0585)	(0.0237)	(0.0496)
After * Insurance * Tobacco Household	-0.129***	-0.116***	-0.113***
	(0.0320)	(0.0440)	(0.0366)
Household Size			0.00382
			(0.00256)
Education (0 = illiteracy, 1 = primary,			0.0510***
2 = secondary, 3 = high school, 4 = college)			(0.0195)
Age			0.00312***
			(0.000371)
No. of Observation	47951	47951	47951
Year Fixed Effects	No	Yes	Yes
R-squared	0.106	0.112	0.141

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.7. Effect of Insurance Provision on Borrowing

VARIABLES	Loan Size (10,000 RMB)		
	(1)	(2)	(3)
After	0.108	1.889***	1.679***
(= 0 if 2000-2002, = 1 if 2003-2008)	(0.0769)	(0.659)	(0.328)
Insurance	-0.111***	-0.111***	-0.133***
(= 0 if control region, = 1 if treatment region)	(0.0109)	(0.0119)	(0.0185)
After * Insurance	0.0481	0.0505	0.0923**
	(0.0769)	(0.0615)	(0.0368)
Tobacco Household	-0.0665**	-0.0669***	-0.147***
(= 0 if No, = 1 if Yes)	(0.0272)	(0.0240)	(0.0172)
After * Tobacco Household	0.0810	0.0636	0.146***
	(0.0583)	(0.0649)	(0.0389)
Tobacco Household * Insurance	0.0441	0.0183	0.0267
	(0.0272)	(0.0309)	(0.0336)
After * Insurance * Tobacco Household	0.0972*	0.134*	0.115**
	(0.0583)	(0.0757)	(0.0556)
Household Size			0.00125
			(0.00656)
Annual Household Income (10,000 RMB)			0.0911***
			(0.0137)
Area of Tobacco Production (mu)			0.00208
			(0.00292)
Education (0 = illiteracy, 1 = primary,			0.0642***
2 = secondary, 3 = high school, 4 = college)			(0.0163)
Age			-0.00103
			(0.000739)
No. of Observation	8,382	8,382	8,382
Year Fixed Effects	No	Yes	Yes
R-squared	0.017	0.029	0.081

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.8. Effect of Insurance Provision on Saving

VARIABLES	Saving Rate (Net Saving Divided by Income)		Net Saving (10,000 RMB)		Ratio of Net Checking to Net Total Saving	
	(1)	(2)	(3)	(4)	(5)	(6)
After	0.0275***	0.0906***	0.0561***	0.190	-0.0212	-0.173***
(= 0 if 2000-2002, = 1 if 2003-2008)	(0.00576)	(0.0311)	(0.00372)	(0.129)	(0.0791)	(0.0545)
Insurance	-0.0166	-0.0149***	-0.00219	0.0101	-0.472***	-0.487***
(= 0 if control region, = 1 if treatment region)	(0.0218)	(0.00391)	(0.00797)	(0.00865)	(0.166)	(0.114)
After * Insurance	0.0504***	0.0294***	0.0834***	0.00267	0.272***	0.269***
	(0.00576)	(0.00509)	(0.00372)	(0.0142)	(0.0791)	(0.0857)
Tobacco Household	0.00781	0.0132***	0.0360	0.0347	-0.352	-0.349
(= 0 if No, = 1 if Yes)	(0.0244)	(0.00263)	(0.0456)	(0.0340)	(0.294)	(0.243)
After * Tobacco Household	0.0211*	-0.00141	0.0562	-0.000966	-0.0420	-0.0534
	(0.0109)	(0.00332)	(0.0583)	(0.0304)	(0.290)	(0.296)
Tobacco Household * Insurance	0.00801	0.00343	-0.00975	-0.0194	0.121	0.120
	(0.0244)	(0.00471)	(0.0456)	(0.0422)	(0.294)	(0.244)
After * Insurance * Tobacco Household	-0.0495***	-0.0124**	-0.0708	0.0414	0.169	0.190
	(0.0109)	(0.00617)	(0.0583)	(0.0359)	(0.290)	(0.299)
Household Size		0.00285***		0.00599*		0.00570
		(0.000544)		(0.00338)		(0.00464)
Annual Household Income (10,000 RMB)		-0.0117***		0.00731		-0.0512
		(0.000448)		(0.0115)		(0.0329)
Education (0 = illiteracy, 1 = primary,		0.0235***		0.0322***		-0.0208
2 = secondary, 3 = high school, 4 = college)		(0.000951)		(0.00796)		(0.0186)
Age		0.000540***		0.000542		-0.00173***
		(8.34e-05)		(0.000429)		(0.000620)
No. of Observation	40,561	40,559	40,561	40,559	20,975	20,975
Year Fixed Effects	No	Yes	No	Yes	No	Yes
R-squared	0.027	0.077	0.012	0.043	0.12	0.13

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.9. Dynamic Effects of Insurance Provision on Borrowing and Saving

VARIABLES	Loan Size (10,000 RMB)	Saving Rate (Net Saving Divided by Income)
	(1)	(2)
Insurance	-0.122***	-0.00960**
(= 0 if control region, = 1 if treatment region)	(0.0136)	(0.00428)
Tobacco Household	-0.0224	0.0127***
(= 0 if No, = 1 if Yes)	(0.0261)	(0.00167)
Tobacco Household * Insurance	0.00750	-0.00168
	(0.0309)	(0.00475)
2001 * Insurance * Tobacco Household	-0.0591	
	(0.0364)	
2002 * Insurance * Tobacco Household	-0.0176	0.00822
	(0.0395)	(0.00860)
2003 * Insurance * Tobacco Household	0.0185	-0.00740
	(0.107)	(0.00919)
2004 * Insurance * Tobacco Household	0.0970	-0.0127
	(0.107)	(0.00903)
2005 * Insurance * Tobacco Household	0.138***	-0.0115
	(0.0395)	(0.0108)
2006 * Insurance * Tobacco Household	0.217***	-0.0187**
	(0.0287)	(0.00955)
2007 * Insurance * Tobacco Household	0.150***	-0.0147**
	(0.0561)	(0.00713)
2008 * Insurance * Tobacco Household	0.210***	-0.0111
	(0.0568)	(0.00902)
No. of Observation	8,382	40,561
Year Fixed Effects	Yes	Yes
Year Dummies * Insurance	Yes	Yes
Year Dummies * Tobacco Household	Yes	Yes
R-squared	0.021	0.057

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.10. Heterogeneity of the Insurance Effect: Production Size and Migration Income

	Production Size			Migration Income		
VARIABLES	Area of Tobacco Production (mu)	Loan Size (10,000 RMB)	Saving Rate (Net Saving Divided by Income)	Area of Tobacco Production (mu)	Loan Size (10,000 RMB)	Saving Rate (Net Saving Divided by Income)
	(1)	(2)	(3)	(4)	(5)	(6)
After	8.180***	1.875***	0.0588***	7.616***	2.018***	0.0655***
(= 0 if 2000-2002, = 1 if 2003-2008)	(3.141)	(0.114)	(0.0137)	(1.807)	(0.0676)	(0.0251)
Insurance	0.764***	-0.0924***	0.0358***	0.381***	-0.104***	-0.00684
(= 0 if control region, = 1 if treatment region)	(0.283)	(0.00290)	(0.00182)	(0.132)	(2.47e-05)	(0.0508)
Tobacco Household		-0.0834***	-0.0172**		-0.102***	-0.00981
(= 0 if No, = 1 if Yes)		(0.00157)	(0.00718)		(0.00645)	(0.0305)
After * Insurance	1.225***	0.0631***	0.0361***	1.493***	-0.0323	0.0318*
	(0.175)	(0.0220)	(0.00197)	(0.0816)	(0.0490)	(0.0187)
After * Tobacco Household		0.113***	0.0222***		-0.00452	0.0247
		(0.00939)	(0.000263)		(0.0288)	(0.0209)
Tobacco Household * Insurance		0.0180***	-0.0233***		0.0437***	0.0263
		(0.00125)	(0.00785)		(0.0101)	(0.0310)
After * Insurance * Tobacco Household		0.0294***	0.00355***		0.258***	-0.0396*
		(0.0108)	(0.000700)		(0.0527)	(0.0206)
Pre-treatment Total Production Area	2.178***	-0.0102***	-0.0288***			
(= 0 if < Median, = 1 if > Median)	(0.637)	(0.00211)	(0.00213)			
Pre-treatment Total Production Area * After	-0.163*	0.203***	0.00760***			
	(0.0873)	(0.0649)	(0.00164)			
Pre-treatment Total Production Area * Insurance	-1.374**	-0.0126***	-0.0327***			
	(0.676)	(0.00234)	(0.00218)			
Pre-treatment Total Production Area * Tobacco Household		0.0295	0.0606***			
		(0.0289)	(0.0195)			
Pre-treatment Total Production Area	0.0177	-0.178***	-0.0187***			
*After * Insurance	(0.112)	(0.0658)	(0.00198)			
Pre-treatment Total Production Area		-0.224***	-0.0309**			
* After * Tobacco Household		(0.0286)	(0.0148)			
Pre-treatment Total Production Area		-0.0107	-0.00221			
* Tobacco Household * Insurance		(0.0289)	(0.0199)			
Pre-treatment Total Production Area		0.291***	-0.00808			
* After * Insurance * Tobacco Household		(0.0290)	(0.0153)			
Pre-treatment Share of Migration Income in Total Income				-0.291**	-0.00283	0.0180
				(0.117)	(0.00211)	(0.0316)
Pre-treatment Share of Migration Income in Total Income				0.239*	-0.146***	-0.00792
* After				(0.124)	(0.00939)	(0.0141)
Pre-treatment Share of Migration Income in Total Income				-1.092***	-0.0279***	-0.0254
* Insurance				(0.0873)	(0.00407)	(0.0315)
Pre-treatment Share of Migration Income in Total Income				-0.280***	0.218***	-0.00658
* Tobacco Households				(0.0948)	(0.0158)	(0.0141)
Pre-treatment Share of Migration Income in Total Income					0.0132	0.0281
*After * Insurance					(0.0608)	(0.0351)
Pre-treatment Share of Migration Income in Total Income					0.194*	-0.0347
* After * Tobacco Household					(0.115)	(0.0236)
Pre-treatment Share of Migration Income in Total Income					-0.0458	-0.0393
* Tobacco Household * Insurance					(0.0768)	(0.0351)
Pre-treatment Share of Migration Income in Total Income					-0.278**	0.0319
* After * Insurance * Tobacco Household					(0.129)	(0.0237)
Observations	34,207	8,382	40,561	34,207	8,382	40,561
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.208	0.030	0.074	0.157	0.036	0.070

Notes: Bootstrap clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.11. Percentage of Households Changing Sector by Region and Year

	Treatment Regions			Control Regions		
Year	Tobacco to Non-Tobacco	No Change	Non-Tobacco to Tobacco	Tobacco to Non-Tobacco	No Change	Non-Tobacco to Tobacco
2000	N/A	N/A	N/A	N/A	N/A	N/A
2001	0	100	0	0.08	99.92	0
2002	0	98.72	1.28	0	94.97	5.03
2003	0.32	99.68	0	0.53	99.42	0.05
2004	0.63	99.37	0	0.32	99.17	0.51
2005	0	99.81	0.19	0	99.63	0.37
2006	0.23	99.53	0.23	0	99.9	0.1
2007	0.4	99.6	0	0.54	99.29	0.17
2008	0.42	99.43	0.14	0.11	99.33	0.56

Bibliography

Ambrus, Attila, Markus Mobius, and Adam Szeidl, "Consumption Risk-sharing in Social Networks," NBER Working Papers 15719, National Bureau of Economic Research, Inc February 2010.

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

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 09 2006, *74* (5), 1403–1417.

Bandiera, Oriana and Imran Rasul, "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, October 2006, *116* (514), 869–902.

Banerjee, Abhijit, G. Chandrasekhar Arun, Esther Duflo, and Matthew O. Jackson, "The Diffusion of Microfinance," *Working Paper*, 2011.

Banerjee, Abhijit V, "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, August 1992, *107* (3), 797–817.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian, "The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States," in "Social Security Policy in a Changing Environment" NBER Chapters, National Bureau of Economic Research, Inc, May 2009, pp. 167–195.

_ , _ , _ , _ , and Katherine L. Milkman, "The Effect of Providing Peer Information on Retirement Savings Decisions," *NBER Working Papers*, August 2011, (17345).

Besley, Timothy and Anne Case, "Diffusion as a Learning Process: Evidence from HYV Cotton," Working Papers 228, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies. May 1994.

Bryan, Gharad, "Ambiguity and Insurance," *Working Paper*, 2010.

Cai, Hongbin, Yuyu Chen, and Hanming Fang, "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, June 2009, *99* (3), 864–82.

Cai, Jing, "Social Networks and the Decision to Insure: Evidence from Randomized Experiments in China," *Working Paper*, 2012.

_ **and Changcheng Song**, "Insurance Take-up in Rural China: Learning from Hypothetical Experience," *Working Paper*, 2011.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From Natural Variation to Optimal Policy? The Lucas Critique Meets Peer Effects," *NBER Working Papers*, March 2011, (16865).

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

_ **, Francisco Galarza, and Stephen Boucher**, "Underwriting area-based yield insurance to crowd-in credit supply and demand," MPRA Paper 24326, University Library of Munich, Germany November 2007.

Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *American Economic Review*, September 2010, *100* (4), 1358–98.

Cole, Shawn Allen, Thomas Sampson, and Bilal Zia, "Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets?," *Journal of Finance*, Forthcoming.

Cole, Shawn, Petia Topalova, , Xavier Gene, Jeremy Tobacman, Robert Townsend, and James Vickery, "Barriers to Household Risk Management: Evidence from India," Working Papers id:4293, eSocialSciences July 2011.

Conley, Timothy G. and Christopher R. Udry, "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, March 2010, *100* (1), 35–69.

Dercon, Stefan and Luc Christiaensen, "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia," *Journal of Development Economics*, November 2011, *96* (2), 159–173.

Duflo, Esther and Emmanuel Saez, "The Role Of Information And Social Interactions In Retirement Plan Decisions: Evidence From A Randomized Experiment," *The Quarterly Journal of Economics*, August 2003, *118* (3), 815–842.

_ , **Michael Kremer, and Jonathan Robinson**, "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya," *American Economic Review*, May 2008, *98* (2), 482–88.

_ , _ , **and** _ , "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, October 2011, *101* (6), 2350–90.

Dupas, Pascaline, "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment," NBER Working Papers 16298, National Bureau of Economic Research, Inc August 2010.

Einav, Liran, Amy Finkelstein, and Jonathan Levin, "Beyond Testing: Empirical Models of Insurance Markets," *Annual Review of Economics*, 09 2010, *2* (1), 311–336.

Ellison, Glenn and Drew Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy*, August 1993, *101* (4), 612–43.

Feder, Gershon, Richard E Just, and David Zilberman, "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, January 1985, *33* (2), 255–98.

Fehr, Ernst and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, March 2007, *97* (1), 298–317.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler, "Testing Enforcement Strategies in the Field: Legal Threat, Moral Appeal and Social Information," *Journal of the European Economic Association*, 2011, *Forthcoming.*

Foster, Andrew D and Mark R Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, December 1995, *103* (6), 1176–1209.

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

Frey, Bruno S. and Stephan Meier, "Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment," *American Economic Review*, December 2004, *94* (5), 1717–1722.

Gaurav, Sarthak, Shawn Cole, and Jeremy Tobacman, "The Randomized Evaluation of Financial Literacy on Rainfall Insurance Take-up in Gujarat," *ILO Microinsurance Innovation Facility Research Paper*, 2011, (1).

Gazzale, Robert S., Julian Jamison, Alexander Karlan, and Dean S. Karlan, "Ambiguous Solicitation: Ambiguous Prescription," Department of Economics Working Papers 2009-02, Department of Economics, Williams College March 2009.

Gin, Xavier and Dean Yang, "Insurance, credit, and technology adoption: Field experimental evidencefrom Malawi," *Journal of Development Economics*, May 2009, *89* (1), 1–11.

Giné, Xavier, Robert Townsend, and James Vickery, "Patterns of Rainfall Insurance Participation in Rural India," *World Bank Economic Review*, October 2008, *22* (3), 539–566.

Granovetter, Mark S., "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78*, 1360–1380.

Haselhuhn, Michael P., Devin G. Pope, and Maurice E. Schweitzer, "Size Matters (and so Does Experience): How Personal Experience with a Fine Influences Behavior," *Working Paper*, 2009.

Holt, Charles A., "Games, and Strategic Behavior: Recipes for Interactive Learning," *Unpublished Manuscript*, 2005.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, "Social Interaction and Stock-Market Participation," *Journal of Finance*, 02 2004, *59* (1), 137–163.

Jackson, Matthew O., *Social and Economic Networks*, Princeton University Press, 2008.

_ , "An Overview of Social Networks and Economic Applications," *The Handbook of Social Economics*, 2010.

_ **and Leeat Yariv**, "Diffusion, Strategic Interaction, and Social Structure," *Handbook of Social Economics*, 2010.

Jensen, Robert, "Agricultural Volatility and Investments in Children," *American Economic Review*, May 2000, *90* (2), 399–404.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl, "Trust and Social Collateral," *The Quarterly Journal of Economics*, August 2009, *124* (3), 1307–1361.

Katz, Elihu and Paul Félix Lazarsfeld, *Personal Influence: the Part Played by People in the Flow of Mass Communications*, Free Press, 1955.

Koher, Hans-Peter, Jere R. Behrman, and Susan Cotts Watkins, "The Density of Social Networks and Fertility Decisions: Evidence from S Nyanza District, Kenya," *Demography*, February 2001, *38* (1), 43–58.

Kremer, Michael and Dan Levy, "Peer Effects and Alcohol Use among College Students," *Journal of Economic Perspectives*, Summer 2008, *22* (3), 189–206.

_ **and Edward Miguel**, "The Illusion of Sustainability," *The Quarterly Journal of Economics*, 08 2007, *122* (3), 1007–1065.

Levitt, Steven D. and John A. List, "What Do Laboratory Experiments Tell Us About the Real World?," *Journal of Economic Perspectives*, 2007, *21* (2), 153–174.

Lusardi, A. and O.S. Mitchell, "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 2007, *54* (1), 205–24.

Lusardi, Annamaria and Peter Tufano, "Debt Literacy, Financial Experiences, and Overindebtedness," March 2009, (14808).

_ **and** _ , "Debt Literacy, Financial Experiences, and Overindebtedness," NBER Working Papers 14808, National Bureau of Economic Research, Inc March 2009.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

Manski, Charles F, "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, July 1993, *60* (3), 531–42.

_ , *Identification Problems in the Social Science*, Cambridge: Harvard University Press, 1995.

Morduch, Jonathan, "Income Smoothing and Consumption Smoothing," *Journal of Economic Perspectives*, Summer 1995, *9* (3), 103–14.

Munshi, Kaivan, "Networks In The Modern Economy: Mexican Migrants In The U.S. Labor Market," *The Quarterly Journal of Economics*, May 2003, *118* (2), 549–599.

_ , "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, February 2004, *73* (1), 185–213.

_ **and Jacques Myaux**, "Social norms and the fertility transition," *Journal of Development Economics*, June 2006, *80* (1), 1–38.

Oster, Emily and Rebecca Thornton, "Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-up," *Journal of the European Economic Association*, Forthcoming.

P.F, B. Berelson H. Gaudet Lazarsfeld, *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*, New York: Columbia University Press, 1944.

Plott, Charles R., *Public Economics, Political Processes and Policy Applications*, Vol. 1, Cheltenham, UK: Edward Elgar Publishing, 2001.

Rogers, Everett M., *Diffusion of Innovations*, 4 ed., Free Press, 1995.

Rosenzweig, Mark R and Kenneth I Wolpin, "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India," *Journal of Political Economy*, April 1993, *101* (2), 223–44.

_ **and Oded Stark**, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, August 1989, *97* (4), 905–26.

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

Song, Changcheng, "Neglect of Compound Interest and Pension Contributions in China," *Working Paper*, 2011.

Stango, Victor and Jonathan Zinman, "Exponential Growth Bias and Household Finance," *Journal of Finance*, December 2009, *64* (6), 2807–2849.

Thaler, Richard H. and Cass R. Sunstein, "Libertarian Paternalism," *American Economic Review*, May 2003, *93* (2), 175–179.

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

Townsend, Robert M, "Risk and Insurance in Village India," *Econometrica*, May 1994, *62* (3), 539–91.

Appendix A

Social Networks and the Decision to Insure

A.1 Supplementary Figures and Tables

Figure A.1. Geographic Location of the Experimental Sites



Figure A.2. The Insurance Indemnity Rule

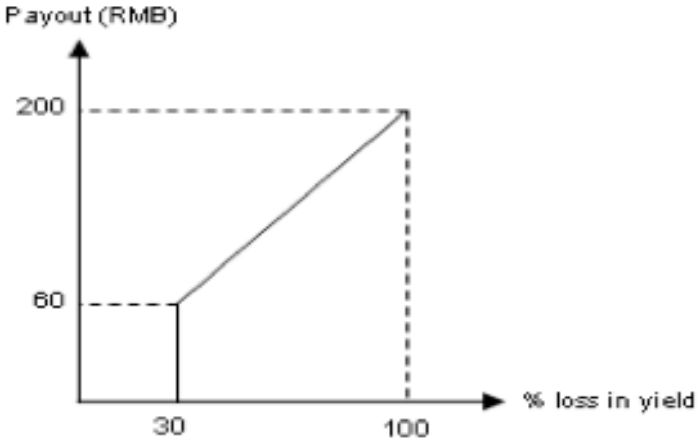


Figure A.3. A Sample Social Network Map Within Beilian Village (ID = 1023)

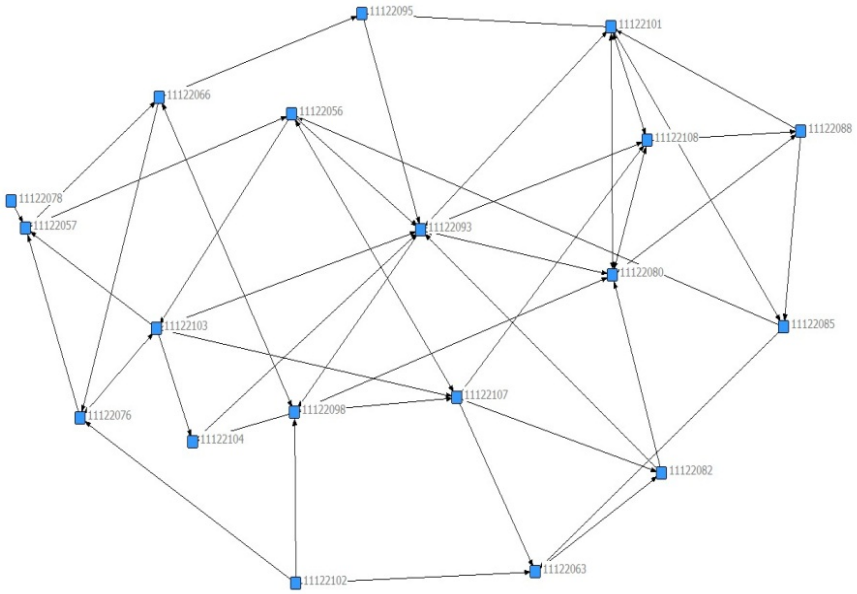


Figure A.4. Probability of Receiving Payouts

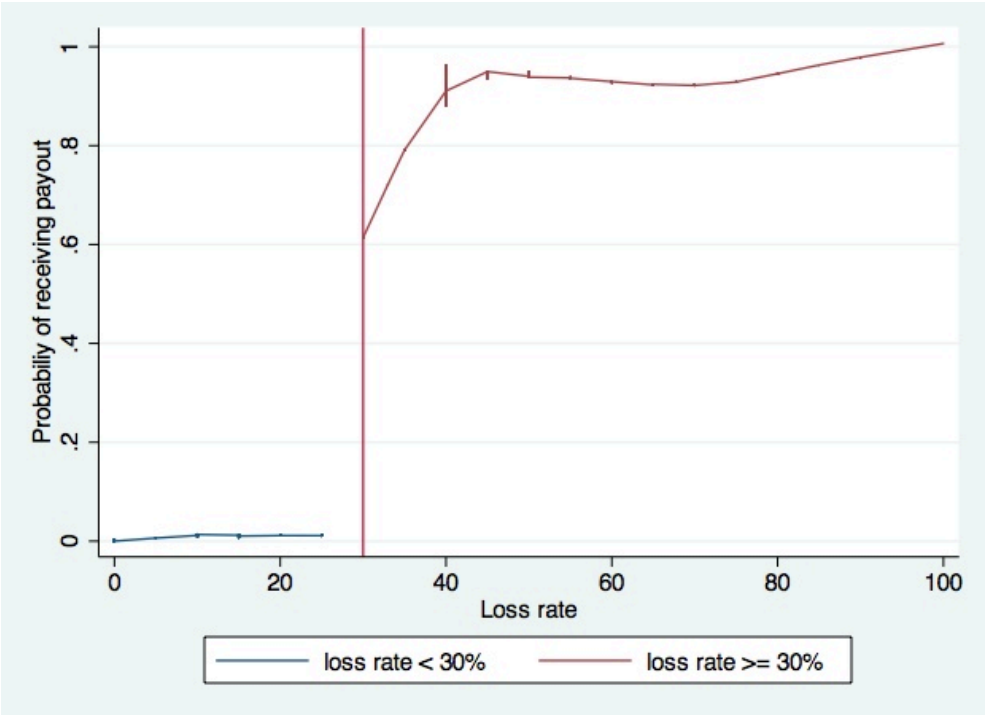


Table A.1. Risk Attitude Questions

	Option A	Option B
1	50 RMB	Toss a coin. If it is heads, you get 200 RMB; if it is tails, you get nothing.
2	80 RMB	Toss a coin. If it is heads, you get 200 RMB; if it is tails, you get nothing.
3	100 RMB	Toss a coin. If it is heads, you get 200 RMB; if it is tails, you get nothing.
4	120 RMB	Toss a coin. If it is heads, you get 200 RMB; if it is tails, you get nothing.
5	150 RMB	Toss a coin. If it is heads, you get 200 RMB; if it is tails, you get nothing.

Note: The risk aversion measure ranges from 0 to 1 and is defined as the fraction of option A choices.

Table A.2 Randomization Check: Session Assignments, Year One

	First Round		Second Round		P-Value
	Simple Session	Intensive Session	Simple Session	Intensive Session	
Gender of Household Head (1 = Male, 0 = Female)	0.908	0.923	0.91	0.915	0.5982
	(0.289)	(0.266)	(0.286)	(0.279)	
Age	51.489	51.091	51.724	51.592	0.6118
	(11.879)	(12.173)	(12.227)	(11.841)	
Household Size	4.902	4.856	4.943	4.945	0.7084
	(2.122)	(2.094)	(2.203)	(2.103)	
Education (0 = illiteracy, 1 = primary, 2 = secondary,	1.193	1.215	1.194	1.17	0.6471
3 = high school, 4 = college)	(0.859)	(0.85)	(0.866)	(0.839)	
Area of Rice Production (mu)	12.965	12.965	11.978	12.247	0.6263
	(15.25)	(26.307)	(14.397)	(21.882)	
Share of Rice Income in Total Income (%)	74.377	74.1	71.887	73.054	0.2812
	(33.878)	(33.553)	(36.015)	(35.414)	
Any Disasters Happened Last Year (1 = Yes, 0 = No)	0.624	0.633	0.634	0.632	0.9627
	(0.485)	(0.482)	(0.482)	(0.483)	
Loss in Yield Last Year (%)	27.042	27.683	27.601	27.651	0.9208
	(18.498)	(18.116)	(18.374)	(17.861)	
Number of Households	1079	1096	1587	1570	

Note: This table checks validity of 1st year session randomization. Standard deviations are in parentheses. P-values reported are F-test of equal means of the four session groups. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3 Randomization Check: Price, Year One and Year Two

	Year One			Year Two		
	OLS Coeff on Price	OLS Coeff on Price Squared	P-Value Joint Test (Price and Price Squared)	OLS Coeff on Price	OLS Coeff on Price Squared	P-Value Joint Test (Price and Price Squared)
	(1)	(2)	(3)	(4)	(5)	(6)
Gender of Household Head	0.007	0.001	0.4374	-0.0006	-0.00004	0.7919
(1 = Male, 0 = Female)	(0.072)	(0.008)		(0.011)	(0.001)	
Age	-0.331	0.096	0.2317	-0.919	0.092	0.3993
	(1.961)	(0.209)		(0.898)	(0.107)	
Household Size	0.105	-0.013	0.8798	-0.099	0.012	0.8028
	(0.236)	(0.026)		(0.228)	(0.023)	
Literacy	0.0113	-0.001	0.9845	0.036	-0.004	0.3782
(1 = Yes, 0 = No)	(0.07)	(0.008)		(0.03)	(0.003)	
Area of Rice Production (mu)	1.085	-0.123	0.7783	1.176	-0.138	0.3922
	(1.574)	(0.185)		(1.199)	(0.118)	
Number of Households	431			1871		

Note: This table checks validity of price randomization in both year one and year two. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4. Effect of Social Networks On Insurance Take-up, Year One: Excluding Households with Less Than Five Friends

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)	
	(1)	(2)
%Network Receiving 1st Round Financial Education	0.324***	0.335***
	(0.0876)	(0.0841)
Intensive Financial Education Session		-0.00444
(1 = Yes, 0 = No)		(0.0337)
Male		0.0124
		(0.0686)
Age		0.00362***
		(0.00127)
Household Size		-0.00970
		(0.00704)
Rice Production Area (mu)		0.00319**
		(0.00124)
Literacy (1 = Yes, 0 = No)		0.0840**
		(0.0332)
Risk Aversion		0.122**
(0-1, 0 as risk loving and 1 as risk averse)		(0.0510)
Perceived Probability of Disaster		0.00213**
		(0.000839)
No. of Observation	1,222	1,204
Village Fixed Effects	Yes	Yes
R-squared	0.089	0.114

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not received 1st round take-up information from us. Households who listed less than five friends are exluded from the estimation. Social network is measured by the fraction of five friends that a household listed who were assigned to a first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table A.5. Determinants of Household Level Social Network Characteristics, Year One

VARIABLES	In-degree	Out-Path Length	In-Path Length	Eigenvector Centrality
	(1)	(2)	(3)	(4)
Male	0.457***	-0.336***	-0.248	0.0179***
(1 = Yes, 0 = No)	(0.146)	(0.120)	(0.152)	(0.00558)
Age	-0.00650*	-0.00133	0.00559*	-0.000313**
	(0.00373)	(0.00355)	(0.00304)	(0.000142)
Household Size	0.0284*	-0.0406***	0.00761	0.00195***
	(0.0151)	(0.0131)	(0.0165)	(0.000624)
Rice Production Area (mu)	0.00961**	-0.00324	0.00882***	0.000306***
	(0.00435)	(0.00287)	(0.00273)	(0.000108)
Literacy (1 = Yes, 0 = No)	0.372***	-0.173**	0.0489	0.00668*
	(0.0937)	(0.0869)	(0.102)	(0.00371)
Observations	4,811	4,245	4,811	4,811
Village Fixed Effects	Yes	Yes	Yes	Yes
R-Squared	0.050	0.176	0.436	0.132

Notes: Robust clustered standard errors in parentheses. Definitions of social network characteristics are in appendix. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

Table A.6. Heterogeneity on the Social Network Effect: Household Characteristics, Year One

VARIABLES	Insurance Take-up (1 = Yes, 0 = No)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
%Network Receiving 1st Round Financial Education	0.435***	0.888*	0.242**	0.0437	0.226**	0.0138	0.369***	0.764***
	(0.113)	(0.513)	(0.107)	(0.238)	(0.0954)	(0.125)	(0.0853)	(0.135)
Average Age	-0.000859	0.00150						
(friends in 1st round financial education)	(0.000831)	(0.00103)						
Average Age		-0.0269***						
*%Network Receiving 1st Round Financial Education		(0.00782)						
Age (own)	0.00369***	0.00231	0.00384***	0.00384***	0.00389***	0.00388***	0.00367***	0.00379***
	(0.00123)	(0.00166)	(0.00122)	(0.00123)	(0.00121)	(0.00124)	(0.00123)	(0.00121)
Age (own)		0.00955						
*%Network Receiving 1st Round Financial Education		(0.00659)						
Average Household Size			0.0105	0.00606				
(friends in 1st round financial education)			(0.00772)	(0.0104)				
Average Household Size				0.0357				
*%Network Receiving 1st Round Financial Education				(0.0543)				
Household Size (own)	-0.00846	-0.00840	-0.00902	-0.0125	-0.00904	-0.00867	-0.00875	-0.00980
	(0.00674)	(0.00662)	(0.00676)	(0.00862)	(0.00674)	(0.00680)	(0.00677)	(0.00667)
Household Size (own)				0.0181				
*%Network Receiving 1st Round Financial Education				(0.0341)				
Average Education					0.167**	0.00351		
(friends in 1st round financial education)					(0.0752)	(0.107)		
Average Education						0.673*		
*%Network Receiving 1st Round Financial Education						(0.350)		
Education (own)	0.0818**	0.0751**	0.0867***	0.0864***	0.0838**	0.0593	0.0847***	0.0877***
	(0.0317)	(0.0313)	(0.0320)	(0.0320)	(0.0323)	(0.0378)	(0.0319)	(0.0316)
Education (own)						0.376		
*%Network Receiving 1st Round Financial Education						(0.283)		
Average Rice Production Area (mu)							-0.000621	0.00420**
(friends in 1st round financial education)							(0.00118)	(0.00212)
Average Rice Production Area (mu)								-0.0242**
*%Network Receiving 1st Round Financial Education								(0.0100)
Rice Production Area (own, mu)	0.00326***	0.00348***	0.00325***	0.00332***	0.00323***	0.00321***	0.00322***	0.00616***
	(0.00114)	(0.00108)	(0.00115)	(0.00115)	(0.00115)	(0.00113)	(0.00114)	(0.00145)
Rice Production Area (own, mu)								-0.0149**
*%Network Receiving 1st Round Financial Education								(0.00704)
Male (1= Yes, 0 = No)	0.0404	0.0458	0.0330	0.0321	0.0385	0.0357	0.0384	0.0439
	(0.0673)	(0.0680)	(0.0670)	(0.0672)	(0.0672)	(0.0673)	(0.0671)	(0.0650)
Risk Aversion	0.117**	0.116**	0.121**	0.122**	0.123**	0.122**	0.120**	0.130***
(0-1, 0 as risk loving and 1 as risk averse)	(0.0494)	(0.0491)	(0.0500)	(0.0500)	(0.0495)	(0.0496)	(0.0492)	(0.0494)
Perceived Probability of Disaster	0.00212**	0.00192**	0.00206**	0.00205**	0.00195**	0.00179**	0.00206**	0.00204**
	(0.000821)	(0.000820)	(0.000820)	(0.000818)	(0.000821)	(0.000839)	(0.000819)	(0.000803)
Intensive Financial Education Session	0.00705	0.0105	0.00593	0.00546	0.00756	0.00998	0.00743	0.00868
	(0.0330)	(0.0326)	(0.0329)	(0.0330)	(0.0328)	(0.0329)	(0.0328)	(0.0326)
Observations	1,255	1,255	1,255	1,255	1,255	1,255	1,255	1,255
Village Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.113	0.125	0.113	0.114	0.116	0.120	0.112	0.112
P-Value of joint-significance:								
%Friends in 1st round financial education		0.0000***		0.1131		0.0069***		0.0000***
Characteristics (friends in 1st round financial education)		0.0026***		0.3077		0.0202**		0.0562*
Characteristics (own)		0.0014***		0.31		0.0143**		0.0001***

Notes: Robust clustered standard errors in parentheses. Results in this table are based on the sample of participants in 2nd round sessions who did not received 1st round take-up information from us. Social network is measured by the fraction of five friends that a household listed who were assigned to a first round intensive session. Village dummies are included in all estimations. *** p<0.01, ** p<0.05, * p<0.1

A.2 Glossary of Network Terminology

In this section, I provide brief definitions of measures of social network characteristics studied in Section 1.4.1.2. More detailed descriptions can be found in Jackson (2008).

- **Clustering Coefficient:**

  Village level. It equals the fraction of pairs of a household's neighbors that are neighbor of each other. This is a measure of how interwoven a household's neighborhood is.

- **Transitivity:**

  Village level. It is defined as the fraction of transitive triads (A-B, B-C, A-C) in total number of triads.

- **Reciprocity:**

  Village level. It equals the fraction of mutually linked pairs in the total number of pairs. There is an equilibrium tendency toward dyadic relationships to be either null or reciprocated, and that asymmetric ties may be unstable. A network that has a predominance of null or reciprocated ties over asymmetric connections may be a more equal or stable network than one with a predominance of asymmetric connections (more of a hierarchy).

- **Fraction of nodes in the giant component:**

  Village level. The share of households in the graph that are in the largest connected component. This is a measure of how interwoven the underlying network is.

- **Degree:**

  Household level.

  1. Out-degree: The number of ties from a household to other households. Households who have unusually high out-degree are those who are able to exchange with many others, or make many others aware of their views (more influential actors);
  2. In-degree: The number of ties received by each household. If a household receives many ties (many others seek advice from them), they are often said to be prominent, or to have high prestige.

- **Average path length:**

  Household level. The average out-path (in-path) length equals the mean of the shortest path to (from) a household from (to) any other households. Shorter out-path (in-path) length means information has to travel less to reach (from) other households from (to) a household.

- **Eigenvector centrality:**

  A recursively defined notion of importance. A household's importance is defined to be proportional to the sum of its neighbors' importance. It corresponds to the ith entry of the eigenvector corresponding to the maximal eigenvalue of the adjacency matrix. This is a measure of how important a node is, in the sense of information flow.

A.3 An Insurance Demand Model

In this section, I present an insurance demand model to explain why social networks can influence both level and slope of the insurance demand curve.

A.3.1 Individual Insurance Demand

A rural household i with wealth ω faces uncertainty about future production income due to possible natural disasters, which will cost him Z. Z is a random variable and follows a normal distribution $\mathcal{N}(\mu_z, \sigma_z^2)$. An insurance product can be purchased to hedge the risk at a premium P. However, due to unfamiliarity with the insurance program, each household has its own perception of the insurance benefit, which is denoted by $\epsilon_i \sim \mathcal{N}(\mu_{\epsilon_i}, \sigma_{\epsilon_i})$. Without insurance contract, the expected utility of the household is

$$\mathbb{E}\left(U(\omega - Z)\right)$$

If the household purchased the insurance contract, then its expected utility is

$$\mathbb{E}\left(U(\omega - P + \epsilon_i)\right)$$

Therefore, the household should purchase the insurance if and only if

$$\mathbb{E}\left(U(\omega - P + \epsilon_i)\right) \geq \mathbb{E}\left(U(\omega - Z)\right) \tag{A.1}$$

Assume that the household has a CARA utility function $U(X) = -e^{-AX}$, then

$$\mathbb{E}\left(U(\omega - Z)\right) = -e^{-A_i(\omega - \mu_z) + \frac{1}{2}A_i^2\sigma_z^2}$$
$$\mathbb{E}\left(U(\omega - P + \epsilon_i)\right) = -e^{-A_i(\omega - P + \mu_{\epsilon_i}) + \frac{1}{2}A_i^2\sigma_{\epsilon_i}^2}$$

Plugging these back into condition (A.1), we have

$$
\begin{aligned}
& -e^{-A_i(\omega-P+\mu_{\epsilon_i})+\frac{1}{2}A_i^2\sigma^2_{\epsilon_i i}} \geq -e^{-A_i(\omega-\mu_z)+\frac{1}{2}A_i^2\sigma_z^2} \\
\iff & \omega - P + \mu_{\epsilon_i} - \frac{1}{2}A_i\sigma^2_{\epsilon_i} \geq \omega - \mu_z - \frac{1}{2}A_i\sigma_z^2 \\
\iff & P \leq \mu_z + \mu_{\epsilon_i} + \frac{1}{2}A_i\left(\sigma_z^2 - \sigma^2_{\epsilon_i}\right) \qquad \text{(A.2)} \\
\iff & \mu_{\epsilon_i} \geq P - \mu_z - \frac{1}{2}A_i(\sigma_z^2 - \sigma^2_{\epsilon_i}) \qquad \text{(A.3)}
\end{aligned}
$$

As a result, on the individual level, households with a higher expectation and a lower uncertainty of the value of the insurance product are more likely to buy it. Since receiving insurance knowledge through either formal financial education or informal information diffusion, or observing friends purchasing insurance or receiving payouts, can all influence households' expectation of the product benefits and the uncertainty about it, I expect that these factors have significant effect on individual insurance demand. Besides, individuals who are more risk averse are more likely to buy the insurance.

A.3.2 Aggregate Insurance Demand

To study determinants of the level and slope of the insurance demand curve, I assume that the perceived benefit of the insurance, μ_{ϵ_i}, is distributed with some CDF $F(.)$ and that the risk aversion coefficient and the variance is the same for all household, $A_i = A, \sigma^2_{\epsilon_i} = \sigma^2_\epsilon, \forall i$. Based on those assumptions, I can aggregate (A.3) to obtain the insurance demand curve:

$$Q(P) = 1 - F\left(P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2)\right) \qquad \text{(A.4)}$$

and the slope of the demand curve

$$\frac{\partial Q}{\partial P} = -f\left(P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2)\right) \qquad \text{(A.5)}$$

where $f(.)$ is the pdf. From equation (A.5), the perceived product benefits, the uncertainty about insurance benefits, and the dispersion on the valuation of the product, could potentially affect the slope of the demand curve.

To give a specific example, let's look at Figure A.5. f_l denotes the original distribution of the perceived expected value of the insurance contract in the population, with a corresponding demand curve D_l in Figure A.6. For people who had more friends exposed to financial education or who received payouts, the distribution changes. First, these people may have higher perceived expected insurance benefits on average. Second, the distribution becomes more concentrated, i.e. smaller variance than before. In Figure A.5, the distribution now

shifts to f_h. As a result, the demand curve will shift upward. In the low price region, because the density of the pivotal value μ_{ϵ_i} is lower, the demand curve will be flatter, as indicated in the shaded region of Figure A.6. The demand reduces sharply over the price region where the corresponding pivotal value of μ_{ϵ_i} has high density, i.e. the concentrated region of the distribution f_h. In year two, during the household survey, we asked households' willingness to pay for the insurance. Using that as a proxy for the expected benefits of the product, I draw the distribution of it and compare the distribution between group of households who have a high proportion of friends receiving payouts (higher than the sample median) and those who have a low proportion (lower than the sample median) in Figure A.7. It shows that the distribution of WTP of households who have more friends receiving payout has a higher pivotal value and is less dispersed.

In order to derive the impact on the insurance demand curve of perceived benefits, dispersion on the product valuation, and the uncertainty about the benefits, I need to specify the distribution of μ_{ϵ_i}. Let $F(.)$ be the CDF of a Normal distribution with mean η and variance ψ^2, and $\Phi(.)/\phi(.)$ be the CDF/PDF of a standard normal distribution. Then $F(x) = \Phi\left(\frac{x-\eta}{\psi}\right)$ and $f(x) = \frac{1}{\psi}\phi\left(\frac{x-\eta}{\psi}\right)$. The demand curve in equation (A.4) becomes

$$Q(P) = 1 - \Phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.6}$$

and the slope of the demand curve is

$$S(P) \equiv \frac{\partial Q}{\partial P} = -\frac{1}{\psi}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.7}$$

- Mean of perceived insurance benefit (η):

$$\frac{\partial Q}{\partial \eta}(P) = \frac{1}{\psi}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.8}$$

$$\frac{\partial S}{\partial \eta}(P) = -\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi^3}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.9}$$

From equation (A.8) and (A.9), an increase in η has a positive level effect on the insurance demand curve, as $\phi(.)$ is positive everywhere. The impact on the slope of demand curve is more subtle. The slope will increase (demand curve will be flatter) if $P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta < 0$, and the slope will decrease (demand curve will be steeper) if $P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta > 0$.

- Dispersion of benefits valuation (ψ):

$$\frac{\partial Q}{\partial \psi}(P) = \frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi^2}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.10}$$

$$\begin{aligned}\frac{\partial S}{\partial \psi}(P) =& \frac{1}{\psi^2}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \\ & - \frac{(P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta)^2}{\psi^4}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \\ =& \frac{\psi^2 - (P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta)^2}{\psi^4}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right)\end{aligned} \tag{A.11}$$

From equation (A.10) and (A.11), an increase in ψ has a level effect on the demand curve. The direction depends on the sign of $P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta$: positive if $P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta > 0$, negative if $P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta < 0$. The impact on the slope of the demand curve depends on the sign of $\psi^2-(P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta)^2$. The slope will decrease (demand curve will be steeper) if $\psi^2-(P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta)^2 < 0$, and the slope will increase (demand curve will be flatter) if $\psi^2-(P-\mu_z-\frac{1}{2}A(\sigma_z^2-\sigma_\epsilon^2)-\eta)^2 > 0$.

- Uncertainty about insurance benefits (σ_ϵ^2):

$$\frac{\partial Q}{\partial \sigma_\epsilon^2}(P) = -\frac{A}{2\psi}\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.12}$$

$$\frac{\partial S}{\partial \sigma_\epsilon^2}(P) = \frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi^3}2A\phi\left(\frac{P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta}{\psi}\right) \tag{A.13}$$

From (A.12) and (A.13), the uncertainty about insurance benefits has a negative effect on the level of demand curve. However, the impact on the slope of demand curve depends on the sign of $P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta$. The impact is positive if $P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta > 0$, and it is negative if $P - \mu_z - \frac{1}{2}A(\sigma_z^2 - \sigma_\epsilon^2) - \eta < 0$.

Figure A.5. An Example of the Distribution of the Perceived Insurance Benefits

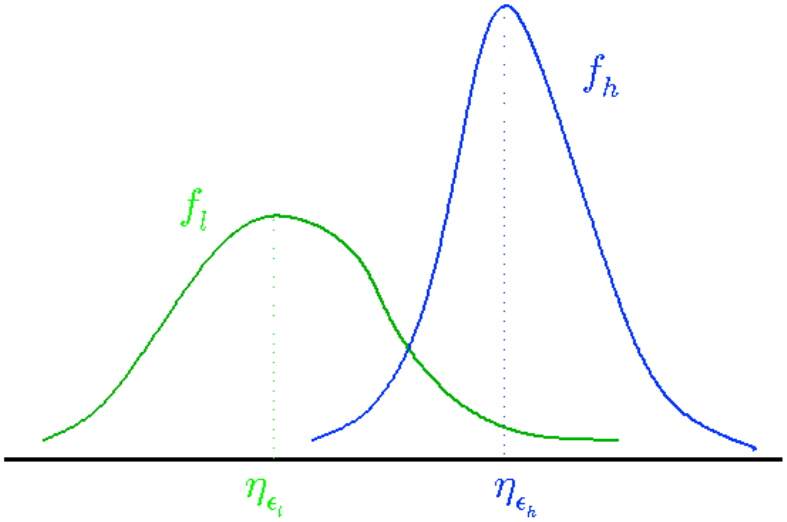


Figure A.6. An Example of Insurance Demand Curve

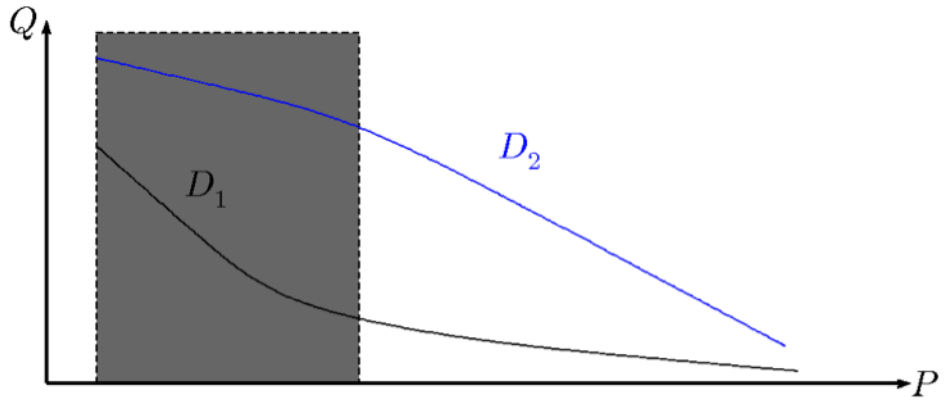
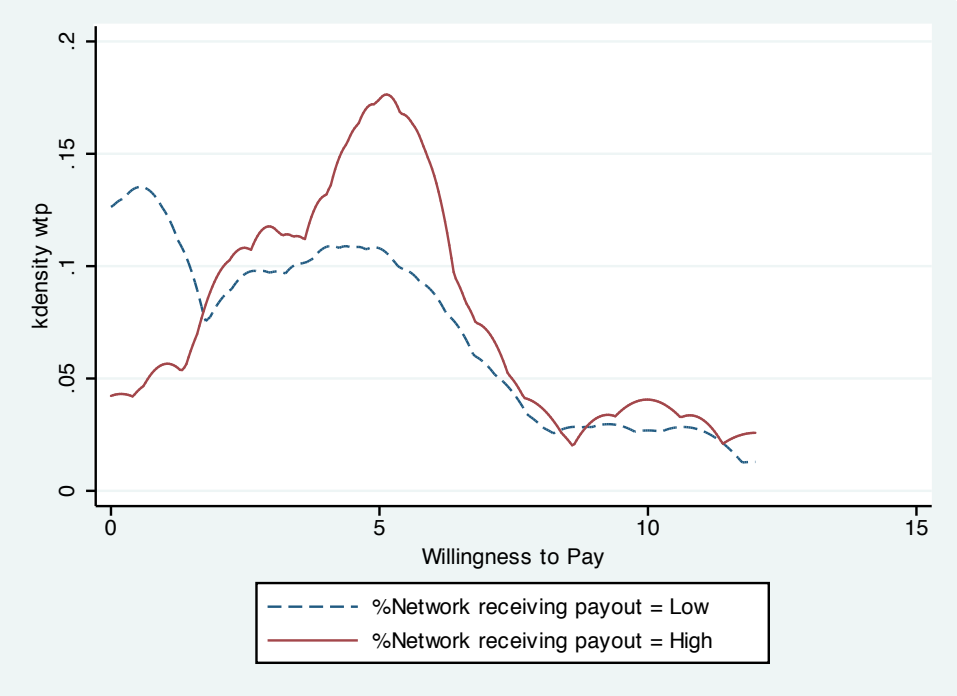


Figure A.7. Distribution of Willingness To Pay for Insurance, Year Two



Appendix B

Insurance Take-up in Rural China

B.1 Simulation Methods

B.1.1 Simulation of Insurance Take-up under Standard Model

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

1. Take a uniform draw of α 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 α, 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.

B.1.2 MLE Estimation of Weight Parameters

We estimate μ_1 and μ_2 with MLE and simulation with the following steps:

1. Take a uniform draw of α from the interval according to each household's choices of riskless options
2. Constrain α to be the draw value and p to be the perceived probability of future disasters from our survey data, then estimate μ_1 and μ_2 with MLE
3. Redo step 1 and 2 for 100 times to generate 100 μ_1 and μ_2 and compare the distribution of μ_1 and μ_2

Table B.1. Risk Attitude Qustions

	Option A	Option B	Choice
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.	

Notes: Risk attitudes were elicited for all the households with questions in table 1. For those who were assigned to play games, risk attitudes were elicited after playing insurance games. 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.

Table B.2. The Effect of Hypothetical Games on Actual Insurance Take-up

Specification:	Logistic Regression			
Dep. Var.:	Individual Adoption of Insurance			
Sample:	All Sample			
	(1)	(2)	(3)	(4)
Game 20%	0.108	0.119	-0.056	0.107
	(0.035)***	(0.034)***	-0.159	(0.036)***
Game 10%	0.045	0.120	-0.097	0.036
	(0.066)	(0.086)	(0.182)	(0.067)
Calculation	0.020	0.012	-0.008	0.020
	(0.045)	(0.048)	(0.189)	(0.043)
Wald test: $\beta_{20}=\beta_{10}$				
p-value	0.2548	0.9956	0.5877	0.2051
Obs.	816	674	132	816
Omitted Treatment	Control			
Mean of Dep. Var. for Omitted Treatment:	0.198			
Social-economic Variables	Y	Y	Y	Y
Fixed Effects for Village and Enumerator	Y	Y	Y	Y
Log Likelihood	-430	-335	-86	-425
Pseudo R-square	0.0933	0.1507	0.0329	0.1045

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.

Table B.3. the Effect of Hypothetical Games on Actual Insurance Take-up

Specification:	Logistic Regression							
Dep. Var.:	Individual Adoption of Insurance							
Sample:	Game				Game with no Information			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Hypothetical Disasters	0.072				0.083	-0.047	0.133	
	(0.037)*				(0.043)*	(0.038)	(0.067)**	
Number of Hypothetical Disasters in First Half of Game (2011-2015)		-0.026	0.107			0.085		
		(0.033)	(0.058)*			(0.038)**		
Number of Hypothetical Disasters in Second Half of Game (2016-2020)		0.075						
		(0.034)**						
Interaction Term (dd5)			-0.051				-0.075	
			(0.050)				(0.057)	
Hypothetical Disaster in 2011				-0.134				-0.140
				(0.082)				(0.105)
Hypothetical Disaster in 2012				-0.134				-0.155
				(0.048)***				(0.052)***
Hypothetical Disaster in 2013				-0.025				-0.030
				(0.071)				(0.080)
Hypothetical Disaster in 2014				0.004				-0.113
				(0.056)				(0.043)**
Hypothetical Disaster in 2015				0.073				0.096
				(0.058)				(0.057)*
Hypothetical Disaster in 2016				0.155				0.164
				(0.039)***				(0.044)***
Hypothetical Disaster in 2017				-0.053				-0.050
				(0.069)				(0.082)
Hypothetical Disaster in 2018				0.120				0.147
				(0.064)*				(0.081)*
Hypothetical Disaster in 2019				0.016				0.037
				(0.079)				(0.060)
Hypothetical Disaster in 2020				0.141				0.160
				(0.063)**				(0.064)**
Obs.	375	375	375	375	292	292	292	292
Omitted Treatment	Control							
Mean of Dep. Var. for Omitted Treatment:	0.198							
Social-economic Variables	Y	Y	Y	Y	Y	Y	Y	Y
Fixed Effects for Village and Enumerator	Y	Y	Y	Y	Y	Y	Y	Y
Log Likelihood	-214	-213	-213	-206	-160	-158	-158	-150
Pseudo R-square	0.0934	0.0966	0.0979	0.1251	0.1219	0.129	0.1307	0.1743

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.

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

Specification:	Logistic Regression								
Dep. Var.:	Individual Adoption of Insurance								
Sample:	Game								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hypothetical Disaster in 2011	-0.134								
	(0.079)*								
# of Hypothetical Disasters (2012-2020)	0.032								
	(0.012)***								
# of Hypothetical Disasters (2011-2012)		-0.102							
		(0.036)***							
# of Hypothetical Disasters (2013-2020)		0.077							
		(0.026)***							
# of Hypothetical Disasters (2011-2013)			-0.077						
			(0.034)**						
# of Hypothetical Disasters (2014-2020)			0.080						
			(0.027)***						
# of Hypothetical Disasters (2011-2014)				-0.053					
				(0.028)*					
# of Hypothetical Disasters (2015-2020)				0.084					
				(0.026)***					
# of Hypothetical Disasters (2011-2015)					0.020				
					(0.033)				
# of Hypothetical Disasters (2016-2020)					0.067				
					(0.040)*				
# of Hypothetical Disasters (2011-2016)						0.013			
						(0.028)			
# of Hypothetical Disasters (2017-2020)						0.039			
						(0.044)			
# of Hypothetical Disasters (2011-2017)							0.005		
							(0.026)		
# of Hypothetical Disasters (2018-2020)							0.068		
							(0.043)		
# of Hypothetical Disasters (2011-2018)								0.025	
								(0.027)	
# of Hypothetical Disasters (2019-2020)								0.057	
								(0.057)	
# of Hypothetical Disasters (2011-2019)									0.004
									(0.014)
# of Hypothetical Disasters (2020)									0.1108
									(0.070)
Obs.	375	375	375	375	375	375	375	375	375
Omitted Treatment				Control					
Mean of Dep. Var. for Omitted Treatment:				0.198					
Social-economic Variables	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fixed effects for village & enumerator	Y	Y	Y	Y	Y	Y	Y	Y	Y
Log Likelihood	-221	-218	-218	-219	-221	-223	-222	-222	-222
Pseudo R-square	0.0628	0.0772	0.0763	0.0733	0.0631	0.0563	0.0601	0.0586	0.0586

Appendix C

The Impact of Insurance Provision

C.1 Two-period model when insurance is not provided

Combine equation (3.1) and (3.4) we can get:

$$U'(C_1) = \beta p U'(C_g)F'(I) = \beta p U'(C_g)(1+R_B) \tag{C.1}$$

$$\Rightarrow \frac{C_g}{C_1} = \frac{F(I)-(1+R_B)B+(1+R_f)S}{C_1} = \beta p(1+R_B)$$

$$\Rightarrow C_1 = \frac{F(I)-(1+R_B)B+(1+R_f)S}{\beta p(1+R_B)}$$

$$= \frac{\left(\frac{1+R_B}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}-(1+R_B)B+(1+R_f)S}{\beta p(1+R_B)} \tag{C.2}$$

Rewrite equation (3.3) as:

$$\beta p U'(C_g)F'(I) = \beta p U'(C_g)(1+R_f) + \beta(1-p)U'\left[(1+R_f)S\right](1+R_f) \tag{C.3}$$

Then combine (C.3) with equation (C.2) we have:

$$\frac{1}{C_1} = \frac{\beta p(1+R_f)}{F(I)-(1+R_B)B+(1+R_f)S} + \frac{\beta(1-p)}{S} = \frac{\beta p(1+R_B)}{\left(\frac{1+R_B}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}-(1+R_B)B+(1+R_f)S}$$

$$\Rightarrow \frac{\beta p(R_B-R_f)}{F(I)-(1+R_B)B+(1+R_f)S} = \frac{\beta(1-p)}{S}$$

$$\Rightarrow \beta p(R_B-R_f)S = \beta(1-p)[F(I)-(1+R_B)B+(1+R_f)S]$$

$$\Rightarrow (1+R_B)B = F(I) - \frac{p}{1-p}(R_B-R_f)S + (1+R_f)S$$

$$\Rightarrow B = \alpha^{-\frac{\alpha}{\alpha-1}}(1+R_B)^{\frac{1}{\alpha-1}} - S\left[\frac{p}{1-p}\frac{R_B-R_f}{R_B+1} - \frac{1+R_f}{1+R_B}\right] \tag{C.4}$$

Plug equation (C.4) into (C.2)

$$\Rightarrow C_1 = \frac{1}{1-p}\frac{R_B-R_f}{\beta(1+R_B)}S \tag{C.5}$$

We know that the total investment is:

$$I = W_0 - C_1 + B - S$$

Replace C_1 and B by (C.5) and (C.4), respectively, we have:

$$\begin{aligned} I = W_0 - \tfrac{1}{1-p}\tfrac{R_B-R_f}{\beta(1+R_B)}S - S + \alpha^{-\frac{\alpha}{\alpha-1}}(1+R_B)^{\frac{1}{\alpha-1}} - S\left[\tfrac{p}{1-p}\tfrac{R_B-R_f}{R_B+1} - \tfrac{1+R_f}{1+R_B}\right] \\ \Rightarrow (1-\alpha^{-1})I = (1-\alpha^{-1})\left(\tfrac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}} \\ = W_0 - \tfrac{1+\beta}{\beta(1-p)}\tfrac{R_B-R_f}{R_B+1}S \\ \Rightarrow S^* = \tfrac{(1+R_B)(1-p)\beta}{(1+\beta)(R_B-R_f)}\left[W_0 + (\alpha^{-1}-1)\left(\tfrac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right] \\ = A * \left[W_0 + (\alpha^{-1}-1)\left(\tfrac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right] \qquad \text{(C.6)} \end{aligned}$$

Now let's consider consumption. Plug the expression of S into equation (C.5):

$$\begin{aligned} C_1 = \tfrac{1}{1-p}\tfrac{R_B-R_f}{\beta(1+R_B)}\tfrac{(1+R_B)(1-p)\beta}{(1+\beta)(R_B-R_f)}\left[W_0 + (\alpha^{-1}-1)\left(\tfrac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right] \\ = \tfrac{1}{1+\beta}\left[W_0 + (\alpha^{-1}-1)\left(\tfrac{1+R_B}{\alpha}\right)^{\frac{1}{\alpha-1}}\right] \qquad \text{(C.7)} \end{aligned}$$

The last variable that we are interested in is the borrowing. According to equation (C.4):

$$\begin{aligned} B = \alpha^{-\frac{\alpha}{\alpha-1}}(1+R_B)^{\frac{1}{\alpha-1}} - S\left[\tfrac{p}{1-p}\tfrac{R_B-R_f}{R_B+1} - \tfrac{1+R_f}{1+R_B}\right] = D + S * E \\ \text{where } D = \alpha^{-\frac{\alpha}{\alpha-1}}(1+R_B)^{\frac{1}{\alpha-1}} \text{ and } E = \tfrac{1+R_f}{1+R_B} - \tfrac{p}{1-p}\tfrac{R_B-R_f}{R_B+1} \end{aligned}$$

C.2 Two-period model when insurance is provided

From equation (3.7), we can see that the expression of optimal investment is:

$$F'(I) = (1+R_B)(1+\delta) \Rightarrow I^* = \left(\tfrac{(1+R_B)(1+\delta)}{\alpha}\right)^{\frac{1}{\alpha-1}}$$

Rewrite equations (3.6) and (3.8) as:

$$\begin{aligned} \tfrac{1}{C_1} = \tfrac{\beta p(1+R_B)}{C_g} + \tfrac{\beta(1-p)\gamma}{C_b(1+\delta)} \qquad \text{(C.8)} \\ \tfrac{\beta p(R_B-R_f)}{C_g} + \tfrac{\beta(1-p)\gamma}{C_b(1+\delta)} = \tfrac{\beta(1-p)(1+R_f)}{C_b} \\ \Rightarrow C_g = AC_b, A = \tfrac{(R_B-R_f)p}{(1-p)\left[(1+R_f)(1+\delta)-\gamma\right]} \qquad \text{(C.9)} \end{aligned}$$

Plug expression (C.9) into (C.8):

$$\frac{1}{C_1} = \frac{\beta p(1+R_B)}{AC_b(1+\delta)} + \frac{\beta(1-p)\gamma}{C_b(1+\delta)}$$
$$\Rightarrow C_b = \frac{\beta p(1+R_B)+\beta(1-p)\gamma A}{A(1+\delta)} C_1 = \frac{\gamma}{1+\delta}(W_0 - C_1 - S + B) - \frac{\gamma}{1+\delta}B + (1+R_f)S$$
$$\Rightarrow \frac{\beta p(1+R_B)+\beta(1-p)\gamma A}{A(1+\delta)} C_1 = \frac{\gamma}{1+\delta}W_0 - \frac{\gamma}{1+\delta}C_1 - \frac{\gamma}{1+\delta}S + (1+R_f)S \quad \text{(C.10)}$$
$$\Rightarrow S = \frac{1}{1+R_f-\gamma/(1+\delta)} \left[\frac{\gamma}{1+\delta} + \frac{\beta p(1+R_B)+\beta(1-p)\gamma A}{A(1+\delta)}\right] C_1 - \frac{\gamma W_0}{(1+R_f)(1+\delta)-\gamma} \quad \text{(C.11)}$$

Combining (C.9) and (C.10) we can get:

$$C_g = [\beta p(1+R_B) + \beta(1-p)\gamma A]\, C_1$$
$$\Rightarrow [\beta p(1+R_B) + \beta(1-p)\gamma A]\, C_1 = f(R_B) - (1+R_B)B + (1+R_f)S$$
$$\Rightarrow B = (1+R_B)^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{\alpha}{\alpha-1}}\alpha^{-\frac{\alpha}{\alpha-1}} - \frac{D}{1+R_B}C_1 + \frac{1+R_f}{1+R_B}S \quad \text{(C.12)}$$

Becasue the total investment is $I = \frac{B+[W_0-C_1-S]}{1+\delta}$, according to equation (C.11) and (C.12) we have:

$$\left(\frac{(1+R_B)(1+\delta)}{\alpha}\right)^{\frac{1}{\alpha-1}} =$$
$$\left[\frac{(R_B-R_f)\gamma+(1+R_B)\big[(1+\delta)(1+R_f)-\gamma\big]}{(1+R_B)(1+\delta)\big[(1+\delta)(1+R_f)-\gamma\big]}W_0 + (1+R_B)^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}}\alpha^{-\frac{\alpha}{\alpha-1}}\right] - [D+E]\, C_1$$
$$\Rightarrow C_1^* = \frac{1}{D+E}\left[\frac{(R_B-R_f)\gamma+(1+R_B)\big[(1+\delta)(1+R_f)-\gamma\big]}{(1+R_B)(1+\delta)\big[(1+\delta)(1+R_f)-\gamma\big]}W_0 + (\alpha^{-1}-1)(\tfrac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}}\right] \quad \text{(C.13)}$$
$$\text{Where } D = \frac{(1+\beta p)(1+R_B)+\beta(1-p)A}{(1+R_B)(1+\delta)}$$
$$E = \frac{R_B-R_f}{(1+R_B)\big[(1+\delta)(1+R_f)-\gamma\big]} \frac{A\gamma+\beta p(1+R_B)+\beta(1-p)A\gamma}{A(1+\delta)}$$
$$\Rightarrow S^* = \frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}\frac{(R_B-R_f)\gamma+(1+R_B)\big[(1+\delta)(1+R_f)-\gamma\big]}{(1+R_B)(1+\delta)\big[(1+\delta)(1+R_f)-\gamma\big]}W_0$$
$$+\frac{(1+R_B)(1+\delta)}{R_B-R_f}\frac{E}{D+E}(\alpha^{-1}-1)(\tfrac{1+R_B}{\alpha})^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{1}{\alpha-1}} - \frac{\gamma W_0}{(1+R_f)(1+\delta)-\gamma} \quad \text{(C.14)}$$
$$B^* = (1+R_B)^{\frac{1}{\alpha-1}}(1+\delta)^{\frac{\alpha}{\alpha-1}}\alpha^{-\frac{\alpha}{\alpha-1}} - \frac{D}{1+R_B}C_1^* + \frac{1+R_f}{1+R_B}S^* \quad \text{(C.15)}$$