Three Essays in Applied Economics

by

Guang Dai

B.E., Harbin Engineering University, 2001
M.A., Nanjing University, 2004
M.A., Université de Toulouse, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2013

Abstract

This dissertation discusses three topics in applied economics.

The first essay examines the causal effect of social capital on individual income by exploiting the historically determined pattern of family name distribution in Chinese villages. Family name distribution impacts social capital through historical inter-lineage rivalry and cooperation. The estimates show a strong first order effect on male villagers, which implies a one standard deviation increase in social capital is equivalent to two to four years of education. No effects on female villagers were found. The gender differentiation could be accounted for by occupation difference: male villagers' income mainly comes from market exchange, while female villagers' income comes mainly from home production. Using a simple model, it is demonstrated that a village's social capital determines its trade scope and therefore income of its residents.

The second essay proposes a general method to identify subjective expectation bias. The method exploits an implication of rational expectations that requires the identical weight of an independent variable in projecting both objective and subjective probabilities. The empirical analysis shows that female seniors do not correctly internalize age information while male seniors fail at internalizing income information. Though cognitive ability and risk aversion can partially explain the results, they are not the sources of the identified biases.

The third essay explores how seniors make long term care insurance (LTCI) decisions by developing a dynamic structural discrete choice model where a rational, risk averse, bequest motivated senior has to decide at each period whether to buy an insurance policy or not. Using the Health and Retirement Survey data, this essay finds substantial heterogeneity in bequest motive that drives LTCI decisions. Specially, the idiosyncratic bequest motive helps to explain why LTCI holders do not experience a higher incidence rate than non-holders.

Preface

This dissertation is original, unpublished, independent work by the author, Guang Dai.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Kevin Milligan for the continuous support of my Ph.D study and research, for his patience, motivation, encouragement, and kindness. His guidance helped me in all the time of research and writing of this dissertation.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Hiro Kasahara and Prof. Marit Rehavi. It is almost impossible without the Hiro's help. Thanks also go to Prof. Siwan Anderson, Prof. Thomas Lemieux, Prof. Thomas Davidoff and Prof. Steven Lehrer for their insightful comments and hard questions.

It was a great pleasure to spend four years with my classmates: Dave Freeman, Xiaodan Gao, Mingzhi Wang, Jinwen Xu, Donna Feir, Mustafa Tugan, Terry Keenland and Andrew Hill. Dave and Xiaodan helped a lot while I was away from the UBC in the past two years.

My colleagues at MSA gave me full support to finish the dissertation. Special thanks go to Harry Chandler, Richard Penn, Mike Nozdryn-Plotnicki, Matt Ayres, Doug Doll and Donna Ehrhardt.

Last, I want to thank my families for their support and love. Thank you, Lucy, for all the time spent together, good or bad.

Dedication

To my wife, Lucy Danxia Mao, and daodaos

Chapter 1

Introduction

This dissertation consists of three essays in applied economics.

The first essay examines the causal effect of social capital on individual income by exploiting the historically determined pattern of family name distribution in Chinese villages. Though social capital has been reported to associate with improved government quality (Putnam, Leonardi, and Nanetti [1993]), better financial development (Guiso, Sapienza and Zingales [2004]) and increased economic growth (Knack and Keefer [1997]) *etc.*, to build a causal argument for social capital upon individual economic achievement faces several challenges, which include measurement and endogeneity problems. This essay employs the rural part of the Chinese General Social Survey to address these problems and exploits a unique feature of the Chinese culture to build causality.

Unlike villages in other societies, a Chinese rural village generally has only a few different surnames among its residents. Families with the same surname are usually from the same ancestor who first settled down at the locality hundreds of years ago. Because of either practical needs or the Confucian philosophy, these families with the same surname were united as a lineage or family clan. In the long run of histo, the interaction between these lineages, which can be rivalry or cooperation depending on resource constraints and objective functions, has a lasting impact on social capital. By instrumenting social capital with historical rivalry or cooperation among lineages, this essay is capable of pinning down the desired causal effect.

The second essay proposes a general method to identify subjective expectation bias. It exploits an implication of rational expectations that requires the identical weight of an independent variable in projecting both objective and subjective probabilities (Manski [2004]). To illustrate, consider the problem of seniors in the Health and Retirement Study (HRS) data set. The seniors were asked to assess their individual probability of entering a nursing home within five years. Consider the information represented by a variable reporting one's health status. Under the rational expectations assumption, the coefficient of health status should be α in projecting the subjective probability, if the coefficient in projecting objective probability were α.

Of course, the objective probability could never be known outside a controlled experimental environment. What is observed is the realization of a random event. Thus, to assess rationality and identify expectation bias, it is necessary to substitute the observed realizations for objective probabilities. The substitution brings an identification issue, which is whether the coefficients are still comparable between the equations of objective probability and of *ex post* realization, by introducing interim events between the formation of expectations and the realization of random events. This essay shows that as long as an orthogonality condition holds, the identification can be successfully achieved.

The last essay explores how seniors shop for long term care insurance (LTCI) policies. It is of importance because of two issues of the market. First, the LTCI market is very thin. Though a senior's lifetime chance of using long term care is more than 40% (Kemper *et al.* [1991]), only about 10% of seniors are covered by any LTCI policies. Second, According to classical asymmetric information theory, the LTCI market is expected to observe a positive correlation between insurance coverage and incident occurrence due to adverse selection and moral hazard. However, recent research noticed that in the LTCI market those covered by an insurance policy do not experience higher incidence rate than those not (Finkelstein and MaGarry [2006]).

To explore underlying preferences that drive the LTCI shopping decisions, this essay develops a dynamic structural discrete choice model where a rational, risk averse, bequest motivated senior has to decide at each period whether to buy an insurance policy or not. In this model, both risk aversion and a bequest motive determine a senior's value function and therefore drive the senior's LTCI decision at each period. By carefully constructing how seniors make insurance and consumption decisions, the chapter is capable of estimating a joint distribution of risk aversion and bequest motives from observed LTCI choices.

Chapter 2

Family Name Distribution: The Effect of Social Capital on Income

2.1 Introduction

Social capital has attracted a great deal of academic attention ever since the work of Putnam, Leonardi, and Nanetti [1993] found a strong correlation between civic engagement and government quality across regions in Italy and concluded that social capital measured by group membership may spur economic success. Subsequently, many have explored the connection between social capital and economic performance. For example, Knack and Keefer [1997] found that a one standard deviation increase in social capital as measured by trust level increases economic growth by more than one half of one standard deviation. But the effect vanishes when social capital is measured by group membership. Laporta *et al.* [1997] reported a similar relationship between the social trust level, judicial efficiency and government corruption. Guiso *et al.* [2004] established that, measured by membership, social capital has a strong effect on financial development in Italy. However, Miguel *et al.* [2005] found no essential relationship between initial social capital and later industrial development across 274 Indonesian districts where a rich set of social capital metrics were employed.

Instead of reporting just another correlation between social capital and economic development, this paper examines the causal effect of social capital on individual economic achievement. By doing so, this paper makes three contributions to the literature. First, by exploiting a cultural institution unique to Chinese rural society, we provide a unique instrument for social capital and thus establish the causal relationship between social capital and individual earnings. Hitherto, the literature was focused on correlation, rather than causal links. Second, we build a simple model to explain how social capital can affect earnings in an economy with imperfect information. Last, our paper presents empirical evidence on different effects of social capital between genders: robust and significant effect of social capital on male peasants but a statistically insignificant effect on females.

The deficiency of empirical work on the causal effect of social capital upon individual economic achievement reflects a number of difficulties. First, social capital is generally measured at the community level. But from an individual's perspective, what is the relevant community and with whom does she or he form a community? Identifying the effect of social capital on individual outcomes requires that each member of the community can actually access to the social capital. Thus, a strong community identification is essential while a segregated community obviously does not meet this condition. Second, social capital, whether measured by average trust level or average club membership of the community members, is almost always constructed endogenously. Attempting to find a valid instrument is notoriously hard, in part due to the lack of sufficient theoretical work on social capital accumulation. Lastly, individuals often self-select their communities based on community attributes, which can make the endogeneity problem more severe.

We use data from the Chinese General Social Survey(CGSS) 2005 to address the above challenges. The CGSS collects very rich information from more than 4000 family heads, which can be either male or female, from 408 Chinese villages. Since a village is an acquaintance society (most villages have hundreds of years of history) in which villagers know each other since birth, almost by definition the residents of a village form a community. The access to village social capital and strong village identification is reinforced by the fact that migration on economic grounds is very weak because of the residence permit *(hukou)* system. During the period studied, nearly the only avenue for inter-village migration was through marriage, and even this was largely confined to women and rarely applied to men. Though rural-urban migration is a challenge to the self-selection issue, we find this migration is mostly based on geographic reasons rather than social capital.

A unique feature of the Chinese culture permits us to instrument the village level social capital by the distribution of family names in a village. Unlike villages in other societies, a Chinese rural village generally only has a few different surnames among its residents. Families with the same surname are from the same ancestor who first settled in a particular locality hundreds of years ago. At least before 1949, when the People's Republic of China was established, those families with the same surname were united as a lineage, or family clan, because of both practical needs and the Confucian philosophy. The long history of a village ensures that the distribution of family names is historically exogenously determined, while the rivalry and cooperation between different lineages have a durable impact on social capital.

To identify the impact of lineages within a village, we utilize two indices constructed from the distribution of family names to measure how rivalry and cooperation affect social capital respectively. The Herfindahl-Hirschman index is used to capture rivalry's lasting influence on social capital, while the polarization index is to seize cooperation's lasting effect. Though inter-village cooperation boosts goodwill, rivalry reduces it. We use these two indices as instruments. The strength of these two instruments is shown by strong first stage relationships. Two robustness analyses provide more confidence in the instrumental variables strategy. To funnel the effect of family name distribution on social capital, lineages must be built based on family names. Anthropology research indicates that such link arises either from the Confucian philosophy or rice production. A significant first stage relationship between the instrumental variables and social capital is not found among these villages where neither the Confucius philosophy is practiced nor rice is cultivated.

OLS estimates show a significant correlation between individual income and social capital as measured by trust. The correlation is robust to controlling for various covariates. A one standard deviation increase in social capital is approximately equivalent to two years of education. However, this significant correlation only exists among male villagers.

To build the case for causality, we use the instruments to estimate the regression and the IV estimates display a fairly large significant first order effect: a one standard deviation increase in social capital is almost equivalent to two to four years of education. Similar to the OLS results, the IV estimation finds no significant effect on female villagers. To correct for a potential weak instrument problem, we estimates robust to weak instrument (RWI) following the method proposed by Chernorzhukov and Hansen [2005]. The estimated RWI intervals show significant effects at the 95% confidence level interval and are robust to various specifications. The RWI estimates also support the absence of any significant effects on females.

We explored the potential mechanism by which social capital affects the income of men but not women. To do this, we build a simple model following Dixit [2003] and Tabellini [2008] where trades are limited by trust or trustworthiness. We assume trade is the only source of income and it happens exclusively between different villages. To build the relationship between trust level and income we utilize the result given by Gleaser [2002] that the answer of a subject to the standard attitudinal survey questions about trust toward strangers predicts the trustworthiness of the individual. In this setup, social capital affects the volume of trade and therefore income in the economy. To understand the gender difference, we explore the difference in occupation between male and female villagers. We suspect that male villagers' income comes principally from free exchanges with markets while females mostly depend on home production. Trustworthiness in a rural economy is crucial for the proper function of markets but has little effect on home production (see Schechter [2007] for an interesting analysis).

Our article is closely related to the contributions by Narayan and Prichett [1999] and Miguel and Gugerty [2004]. Narayan and Prichett studied the role of social capital in determining household income in rural Tanzania by constructing a measure of social capital from a large scale survey data set. They also found a result similar to ours: a one standard deviation increase in social capital increases a household income by at least 20 to 30 percent, equivalent to three years of education. However, their conclusion is vulnerable to the critiques about the validity of the instrumental variables. Their instrument choice is built on the assumption that income is not directly influenced by the trust level in strangers or government officials. Miguel and Gugerty examined ethnic diversity and local public goods in rural western Kenya and found ethnic diversity is associated with lower primary school funding and worse school facilities. They concluded the inability to impose social sanctions in an ethnically diverse community can lead to collective action failures. In this paper, we borrow from their idea about ethnic diversity but instead focus on the lineage diversity and its impact on social capital.

This rest of this article is organized as follows. In section 2.2, we present a simple model to explore why social capital matters and discuss the estimation framework. Section 2.3 describes the data and the measurement of social capital. Estimates from OLS are presented as a benchmark analysis. Section 2.4 is devoted to building the validity of the instruments. We first give a brief introduction to family names in villages and argue for the exogeneity of the surnames distribution using historical and empirical evidences. To prepare for the first stage analysis, we describe in detail the interplay of rivalry and cooperation of the inter-lineage relationship. Two indices are constructed to measure different sides of inter-lineage relationship. First stage analysis shows a strong relationship between the instruments and social capital. Robustness checks provide evidence on the validity of the instruments. Section 2.5 presents our main results and a number of robustness checks and section 2.6 provides conclusion.

2.2 Motivating Theory and Estimation Framework

2.2.1 Motivating Theory

The mechanism through which social capital affects economic performance has been extensively examined by recent literature. However, the link from social capital to individual economic achievement is rarely explored. Narayan and Prichett [1999] is one of few papers that briefly discusses several channels through which social capital would affect individual economic achievement.

In this subsection, we present a simple model that is designed for our empirical analysis. Thus, the model only partially explains how social capital matters at individual level. While it may not wholly explain the link, we believe that the mechanism it posits is particularly interesting.

The central objective is to show how the trust level of a village determines its volume of trade with others and therefore the income of its residents since the model assumes income is exclusively from inter-village exchanges. Bowlus and Sicular [2003] discussed inter-village movement of labour and concluded the factor market was underdeveloped in rural China. Chen, Mu and Ravallion [2009] discussed the spillover effect of inter-village trade. However, this assumption is relatively harmless because the model can be expanded to allow income from intra-village production as well. In a recent paper by Guiso, Sapienza and Zingales [2009], using data on bilateral trust between European countries, these authors found that lower bilateral trust leads to less economic interaction even after controlling for national characteristics.

Throughout we assume a subject's trusting level towards strangers represents the level of trustworthiness of the subject. This is justified by the result in Glaeser *et al.* [2000] where they found the standard attitudinal survey questions about trust predict trustworthy behavior much better than trusting behavior.

Consider the following static model adapted from Dixit [2003] and Tabellini [2008]. Villages indexed by their normalized social capital $k_i = i \in [0, 1]$ are uniformly distributed on interval $[0, 1]$. The uniform distribution assumption is just for convenience and our conclusion has no bearing on it. For any $i \in [0, 1]$, village k_i is resided by a continuum of mass 1 villagers. Among these villagers, fraction k_i are trustworthy denoted as type 1 and the remaining $1 - k_i$ are not trustworthy and denoted as type 0. Villagers know which village other villagers are from but not their types. Thus, when two villagers from two different villages meet, they only know the probability that the other is trustworthy. We assume the villagers' income only come from trades with residents of other villages. For example, consider villagers i_1 and i_2 both from village k_i and villagers j_1 and j_2 both from village k_j. This assumption only allows trade between i and j,but not among is or among js.

To complete the model, we make some assumptions regarding to how trades happen and payoffs from trades. First, villager i from k_i matches with villager j from k_j with equal probability to complete a trade for any j not equal to i. Since there are infinite villages, the probability is zero for any specific i and j. Alternatively, We could assume a locally biased matching probability like Dixit [2003], but this could needlessly complicate our analysis. Second, once i and j match, they simultaneously choose trade or not. If any one of them chooses not to trade, each gets payoff zero and the interaction ends. If both choose to trade, they each get a baseline payoff 1 times the type combination multiplier, which is given by the following matrix:

$$\begin{array}{cc|cc} & & \multicolumn{2}{c}{\text{villager } i} \\ & & \text{type 1} & \text{type 0} \\ \hline \text{villager } j & \text{type 1} & k_i + k_j, k_i + k_j & -(k_i + k_j), \frac{k_i+k_j}{2} \\ & \text{type 0} & \frac{k_i+k_j}{2}, -(k_i + k_j) & \frac{k_i+k_j}{2}, \frac{k_i+k_j}{2} \end{array}$$

where type 1 corresponds to a trustworthy villager and type 0 an untrustworthy one. Notice the structure of the multiplier is very similar to the to payoff structure of the prisoner's dilemma. Specially, if both traders are trustworthy, then each can benefit $k_i + k_j$ from the trade; if both are not trustworthy, each benefits $\frac{k_i+k_j}{2}$; finally, if one is trustworthy and the other not trustworthy, then type 1 loses $k_i + k_j$ and type 0 gets $\frac{k_i+k_j}{2}$. As in any prisoner's dilemma game, we assume throughout that the loss from being cheated is at least as large as the benefit from cheating.

Another feature of the payoff matrix is that the payoff depends on the trustworthiness of k_i and k_j, which can be justified by the assumption of transaction cost reduction or a psychological benefit from trade. This feature is crucial for our result because otherwise only the trustworthiness k_i would affect the trading strategy.

We are especially interested in two questions: Does the expected income of an individual from village k_i, regardless of the trustworthiness type, increase with k_i? If so, which type of villager from the same village k_i has a higher income? To answer these questions, we introduce the following proposition:

Proposition. *The expected income of villager i from village k_i is an increasing function of k_i, the social capital of the village.*

Proof. We consider the Bayesian Nash equilibrium. Obviously, the type 0 villagers always trade and type 1 villagers trade only if her partner from $k_j \geq 0.5$. The expected income of a type 1 villager is $\int_{0.5}^{1}(k_i + k_j)(k_j - (1 - k_j))dk_j = \frac{5}{24} + \frac{k_i}{4}$ for $k_i \geq 0.5$ and zero for $k_i < 0.5$. The expected income of a type 0 villager is $\int_0^{0.5} \frac{k_i+k_j}{2}(1 - k_j)dk_j + \int_{0.5}^{1} \frac{k_i+k_j}{2} dk_j = \frac{11}{48} + \frac{7}{16}k_i$ for $k_i \geq 0.5$ and $\int_0^1 \frac{k_i+k_j}{2}(1 - k_j)dk_j = \frac{5}{12} + \frac{k_i}{4}$ for $k_i < 0.5$. Q.E.D. □

The above proposition clearly shows a mechanism through which social capital can have a positive effect on income. The rest of the paper explores this claim. The proof indicates the income of type 0 villager is larger than the income of a type 1 villager from the same village. [1] Intuitively, a type 0 villager will always be made better off by a trade, and will therefore engage in making transactions whenever encounter with others.

2.2.2 Estimation Framework

As criticized by Durlauf [2001], empirical work on the causal effect of social capital on economic achievement generally suffers from various identification problems. The problems mainly arise from the lack of a solid theory of

[1] Of course, we can endogenize the payoffs for type 1 and type 0 villagers such that both types have the same income in equilibrium. We choose not to do so in order to keep the model simple.

social capital, *e.g.*, definition ambiguity, measurement confusion and formation and function vagueness (see Arrow [1999] and Solow [1999] respectively). Because of the theory deficit, social capital is often measured through the decisions made by individuals. In empirical work, the two most prevalent measurements of social capital are the trust level and membership status in a community. Without a solid theory, it is difficult to claim that there are no unobservables systematically different between those who trust people and those who do not. Similarly, communities with higher social capital may be systematically different than others with regard to unobservable attributes. This becomes more serious if individuals can migrate based on social capital. As we will see later, one of the merits of our data is that inter-village migration is close to impossible due to the strict residence permit system, *hukou*, in China. Technically, without the exchangeability [2] assumption, we cannot exclude the presence of unobserved heterogeneity in the sample under study.

The identification issue is further entangled by reverse causality and neighbour effects. For example, Glaeser *et al.* [2000] reported a strong positive effect of education on trust. The human capital theory has long established the effect of education on income. Using cross country data, Knack and Keefer [1997] showed a strong correlation between economic outcomes and social capitl as measured by trust level. Alesina and Ferrara [2002] reported a robust relation between income and trust level using individual data from the General Social Survey from 1974 to 1994. Since social capital is measured at the group level, it is necessary to disentangle the social capital effect from neighbourhood effects if any. The neighbourhood effects have long been analyzed by Manski [1997] and Brock and Durlauf [2001].

With these considerations, we formulate the following structural equa-

[2] A collection of random variables ϵ_i is exchangeable if for every finite sample of the random variables, $\epsilon_{i_1}...\epsilon_{i_N}$, and every permutation operator $\rho()$, $\mu(\epsilon_{i_1} \leq a_1, ..., \epsilon_{i_N} \leq a_N) = \mu(\epsilon_{\rho(i_1)} \leq a_1, ..., \epsilon_{\rho(i_N)} \leq a_N)$. See Durlauf [2001] for a good discussion on exchangeability.

tions:

$$
\begin{aligned}
INCOME_{i,j} = &\quad aTRUST_{i,j} + bSC_j + cX_{i,j} + d\bar{X}_j + eY_j + \varepsilon_i + \varepsilon_j \quad (2.1a) \\
SC_j = &\quad \gamma_1 Y_j + \gamma_2 Z_j + \upsilon_j \quad (2.1b) \\
TRUST_{i,j} = &\quad \eta_1 SC_j + \eta_2 X_{i,j} + \eta_3 INCOME_{i,j} + \xi_i + \xi_j \quad (2.1c)
\end{aligned}
$$

where i and j represent individual i and community (or village in our paper) j, $INCOME_{ij}$, $TRUST_{ij}$ stands for income of individual ij and the answer to the general trust question reported by individual ij. In the literature of social capital, trust is measured through reports to the question "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". As discussed, the inclusion of $TRUST_{i,j}$ is necessary for capturing the unobservables to ensure the exchangeability assumption. SC_j is the social capital of village j measured by averaging $TRUST_{i,j}$ across i for each j. $X_{i,j}$ is a vector of individual characteristics, $\bar{X}_j$ is the average of $X_{i,j}$ across i within village j. The inclusion of $\bar{X}_j$ is necessary to disentangle the neighbourhood effects. The neighbourhood is usually defined as $\bar{X}_{-i,j}$, the average that excludes the own person's characteristics. We choose $\bar{X}_j$, which does not exclude the own person's characteristics due to sample size consideration. However, main results do not rely on the choice, see Table 2.7. Y_j is a vector of village attributes. Z_j is the excluded instrument for SC_j. The potential endogeneity relations depend on the correlation between ε_i and ξ_i, ε_j and υ_j and ε_i and ξ_j.

The equation 2.1a states the relationship between a villager's income and individual and village attributes. The equation 2.1b states the relationship between social capital and instruments. This equation implicitly assumes that social capital is a characteristic of a village on which no single individual has impact. The equation 2.1c states that individual trust decisions are based on the group trust level (social capital) and individual attributes. Notice that by taking expectation on both sides and using the *iid* assumption, we get what Manski [1993] called the endogenous effect.

We have two alternative specifications for the estimation framework. First, we can estimate equation 2.1a directly, which has the merit that it captures the potential individual unobservables by including covariate $TRUST_{i,j}$ and the demerit that it does not consistently estimate the coefficient of SC_j since from 2.1c it is obvious $TRUST_{i,j}$ absorbs some effect of SC_j. Second, we can use the reduced form as the estimation framework. Combining the equations 2.1a to 2.1c, one therefore arrives at the reduced form expression

$$INCOME_{i,j} = \beta_0 SC_j + \beta_1 X_{i,j} + \beta_2 \bar{X}_j + \beta_4 Y_j + e_i + e_j \tag{2.2}$$

where β_0 is the coefficient of interest. Applying the reduced form to build a causal relation implicitly requires that ξ_i in equation 2.1c satisfies the exchangeability condition. This can be justified by the setup in the model that individual trust decision is based on group attributes. In general, the reduced form is more appropriate since we are mainly focused on the effect of social capital. Thus, in the following regression, we report estimates of β_0 by adopting equating 2.2. We also estimate the coefficient of interest by using equation 2.1a and arrive at similar results.

2.3 Data and OLS Estimates

2.3.1 Data

The main data set in the paper is the rural part of the Chinese General Social Survey 2005 (CGSS) conducted by the department of sociology, Renmin University, China. This data set was collected in October 2005 through interviews with households from 401 villages in 75 counties. About ten families from each village are randomly picked to take the survey. A household head, who can either be the husband or wife, is asked for the household members' information, individual basic information, family background information, value and attitude, and community governance questions. The individual's basic information includes his/her annual income, education, age *etc*; family background includes his/her parents' education.

Among the questions in the value and attitude part is the question like "When it is not directly related with money, to what extent do you think the following people are trustworthy?". The answers are scaled from 1 to 5 where 1 represents "most people can't be trusted", 2 "many can't be trusted", 3 "half can be trusted and half can't be trusted", 4 "many can be trusted" and 5 "most can be trusted". The "following people" ranges from your neighbours, the villagers with the same surname, classmates, and strangers *etc.*. Trust level in strangers is chosen as a measure of social capital. The first reason for this choice is the reliability of this answer. Since many of the interviews are conducted under the presence of neighbours or friends, the answers to friends or neighbours may be biased. Second, our motivating model suggests that the mechanism through which social capital affects income hinges on the private information of a villager's trustworthiness type, which is most suitable if trust attitude is toward strangers. Third, Platteau [2000] stressed that the generalized codes of good conduct like honest and trustworthiness towards a population at large are essential for a modern market economy. Therefore, social capital as measured as trust to strangers can better capture causal effect in a market economy environment. Further, we adopt the mean of the trust answer as the village level social capital following convention. Another measure of social capital, the membership in clubs and associations, is not adopted here because of the extreme low variation in the sample.

Figure 2.1 shows village average trust level in different population groups. From left to right, the adjacent line, whiskers and divided box are corresponding to the lower adjacent value, 25 percentile, median, 75 percentile and upper adjacent value. It illustrates that trusting is most limited to those villagers who have close connections, *i.e.*, people from the same village. The trust in strangers is very low: more than half the villages on average say "many strangers can't be trusted". Notice the differences between the trust in villagers with the same surname and with a different surname, which is significant at 1 percent significance level.

As Deaton [1997] pointed out, it is difficult to accurately measure the income of self-employed, especially peasants. The difficulty is partially because of the sensitivity of the topic, which causes the surveyed income to be underreported or even biased. However, this does not seem to be a se-

Figure 2.1: Trust Level in Different Groups

rious issue in our data set because in the Chinese rural villages, personal income is always a public topic among neighbours, friends, and relatives. Family members who are employed in agriculture or family business makes it hard to distinguish personal and family income. It makes the estimate of income more difficult. Here, the potential difficulty is the inability to measure opportunity cost. This does not apply to our case since the alternative of working in town village enterprises(TVEs) can be used as a benchmark. Based on these observations, we claim the individual income in the data set is reliable. To verify it, we compare the reported income in the data set with the peasants income from Chinese Household Income Project (CHIP) 2002 conducted by National Bureau of Statistics, which is regarded as the most reliable data in income. Following the selection criteria in table 2.1, we find the average (standard deviation) income reported in the CHIP 2002 data set is 8.30(.74) for male and 8.29(.66) for female, which are very close to the values in table 2.1. We also compare the average income using different criteria, and find no material differences.

The data set also includes some village level information from the village head, which includes a village's area, its population, its distance to the county seat and to the nearest market, its literacy rate and high school graduation number, amount of clinics, schools and TVEs (town village enterprises) in the village, and landform of the village. One of the special features of the data set is the inclusion of proportions of families who have one of the top three surnames in the village. For example, in one of the villages, the proportions are 40%, 30% and 20%. Thus, in the village 40% of families have the same surname and this surname is the largest one in the village, and so on. Notice the total proportion of the top three surnames is 90% which implies there are 10% of families in the village with family name(s) other than the top three.

During the study period, several features of the Chinese society are especially notable. To control for rural-urban migration, a residence permit system, *hukou*, was first created in the later part of the 1950s and is still effective today. The key part of rural-urban migration under the *hukou* system is the *hukou* conversion process, which basically requires a rural resident, *e.g.*, a peasant in a village, to meet both qualification and quota criteria to be eligible to migrate to an urban area. It is estimated the annual quotas were extremely small, about 0.01 to 0.02 percent of the rural population of each locale. Since 1980s, many rural labours migrated to urban areas though working on urban jobs and residing for the most part in towns and cities [3], though they are not legally considered urban workers since they have no access to local schools, urban pension plans, public housing and other rights that are available to those with urban *hukou*. This urban-rural migration could potentially raise a self-selection issue. However, as pointed out by Zhao [1999], more people actually choose not to migrate and the migration that does occur is largely circular. Further, such migration is mostly based on geographic choices rather than social capital considerations (see Zhang and Song [2003]).

On the other hand, state control of land closes the channel of rural-rural migration. Even today, private land ownership does not exist. Though a peasant is entitled to have usage right of some arable land in his/her birth village, the right is not transferred to the peasant if the peasant migrates to other villages. As become clearer with our discussion, one possible threat to our adoption of instrumental variable is that those families with a dominate surname are assigned more land or more fertile land even if allocated land size is similar. However, Brandt *etal.* [2002] analyzed a survey data and failed to reject the hypothesis that cultivated land is allocated on average in direct proportion to family size. Therefore, the *hukou* system and absence of land ownership makes the Chinese village an ideal place to study social capital.

In this paper, we only consider villagers aged 20 to 55 with an income greater than one thousand two hundred Chinese Yuan. The age threshold is chosen based on the observation that 20 to 55 is regarded as the work age in many rural areas. The income threshold is to eliminate these not actively participate in labour market. Our empirical results are not sensitive to the choice of the thresholds. Table 2.1 presents the basic descriptive statistics

[3] The number is estimated to be about 190 million in 2009, increasing from about 20 million in 1980s. See Chan and Zhang [1999] and Chan [2010].

on key characteristics. Note that panel B lists the village level averages of age and educations defined as proportions of the villagers satisfying some threshold conditions. In a previous version, these variables were defined as corresponding mathematical averages. In either case, the results are almost identical.

Instead of clustering our data set at the village level, our regression analysis clusters at the county level. Many students have pointed out that the Chinese peasants mostly limit their social and economic activities within a county. Therefore, villagers' income from the same county are highly correlated, which is also reinforced by the fact that the county is the basic administrative unit. Another practical reason is that clustering at the village level sometimes could not generate enough observations for our regressions. We also perform the analysis clustering at the village level where possible and find no substantial difference.

2.3.2 OLS Estimates

Table 2.2 reports OLS estimates of individual income on social capital and other controls. These results are useful both to show correlation in data and for comparison to the the IV estimates. Panel A pertains to the estimates using only male villagers. Column (1) reports the estimate from the basic regression without any other controls. The estimate shows that a one unit increase in social capital is correlated with 15% higher income, a relation significant at the 95% confidence level. Column (2) reports the estimate after controlling for individual characteristics including age, education and parents' education. Column (3) additionally controls for the same characteristics across villages, as defined in as defined in 2.1. As discussed, these variables are used to disentangle any neighbour effects. Column (4) additionally controls for village characteristics including the variables listed in Table 2.1. The results in Panel A present a robust correlation between income and social capital. Panel B displays the results for female villagers. In contrast with the estimates in panel A, the results indicate there is no significant correlation between income and social capital for female villagers

Table 2.1: Descriptive Statistics

	Value	Observations
Panel A: Individual Characteristics		
Income: Male	8.27(.73)	1247
Female	8.08(.64)	1125
Education: Male	6.97(2.89)	1308
Female	5.25(3.35)	1281
Age: Male	39.77(9.29)	1308
Female	38.55(8.98)	1281
Father's education: Male	2.32(2.79)	1308
Female	2.64(3.54)	1281
Mother's education: Male	1.53(1.89)	1308
Female	1.80(2.63)	1281
Panel B: Average Characteristics		
Villager Aged	0.45(.19)	399
Villager Educated	0.56(.22)	399
Father Educated	0.08(.10)	399
Mother Educated	0.02(.06)	399
Panel C: Village Characteristics		
Social Capital	1.87(.59)	399
Population(log)	6.03(.71)	393
Area(km^2)	7.47(.91)	391
Distance to county seat(km)	29.49(21.87)	393
Distance to market	4.89(7.20)	393
Clinics	2.21(.2.26)	292
Primary Schools	1.87(3.61)	295
Illiterate(percentage)	.21(.33)	390
High school graduates	.27(.42)	389
Number of TVEs	2.96(6.65)	389

Notes: Standard deviations are in parentheses. In Panel B, the *villager aged* is the proportion of villagers that are older than 45, and the *villager educated, father educated* and *mother educated* are the proportion of villagers, villagers' fathers and villagers' mothers, respectively, that have at least 5 years of formal school education. Social capital is measured as the village average trust level in strangers.

Table 2.2: Effect of Social Capital on Income: OLS Estimates

	(1)	(2)	(3)	(4)
	Panel A: Male			
Social Capital	.15(.06)**	.14(.05)**	.14(.05)***	.14(.05)***
R^2	.01	.13	.15	.19
F	6.07	15.21	12.41	7.43
Obs.	1247	1156	1134	1113
	Panel B: Female			
Social Capital	.02(.06)	.02(.06)	.01(.05)	.03(.06)
R^2	.00	.05	.09	.10
F	.14	5.71	3.92	5.16
Obs.	1125	1034	1031	1001
Controls:				
Individual Charac.		YES	YES	YES
Average Charac.			YES	YES
Village Charac.				YES

Notes: Standard errors are in parentheses. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is income.

regardless of the specifications. This result is very informative in that it posits a channel to think how social capital affects individual economic outcomes, which is still largely unexplained in academic research. Because of potential endogeneitty problems, we cannot conclude that there is evidence of an causal effect.

2.4 First Stage Analysis

Because of reverse causality and omitted variable problems, OLS estimates are unlikely to uncover the causal effect of social capital on economics achievement. To build the case for causality, we outline a source of exogenous variation in social capital with regards to the economic outcome today. Following the framework, we then formally analyze the first stage relationship. We examine the exclusion assumption in the next section.

2.4.1 Village, Lineage and Family Name

The exogeneity comes from village level variation in family name distribution, which is equivalent to the claim that family name distribution is historically determined hundreds of years ago and has remained essentially unchanged since then. Moreover, the residence permit regulation and one child policy reinforce the claim.

Using historically determined events as excluded instruments is not novel in economics; *i.e.*, Acemoglu and Johnson [2005] applied colonial history to instrument for two different types of institutions. Adopting distribution as an exogenous variation makes our first stage analysis more subtle since the validity of the instruments now relies on appropriate constructions from the distribution. The exclusion assumption is warranted by the fact that lineage system was abolished and replaced with political class by the Communist Party in 1950s. This subsection presents detailed historical and empirical evidence to support our claim after a brief introduction to the background of a typical Chinese village, which is mostly adapted from Hu [1983] and Baker [1979].

Within the structure of Chinese society, lineage occupied a prominent place for many centuries until 1949 when the People's Republic of China was established. A lineage is a corporate group of families which celebrates ritual unity and is based on demonstrated descent from a common ancestor (Watson [1965]). Therefore, a lineage traces its ancestry to one ancestor who first settled in a given locality. The rites in his honor and those for later ancestors serve as a reminder of kinship bonds. However, women were excluded from the lineage since the inheritance of the kinship was predominantly patrilineal.

A lineage differs from a family in its scope and functions. A family, which includes parents, children and grand children, is an economic unit to facilitate child rearing, while a lineage is a patrilineal clan, including all the families with husbands descended from the same distant ancestor. A lineage is a much larger but looser organization compared with a family. Generally, a lineage is cooperative in the sense that member families of a lineage own property (usually land) in common. However, this property is for religious, educational and relief purpose rather than a means of livelihood. Where a lineage constitutes one village, village affairs are managed by its leaders. Furthermore, a lineage is interested in promoting the social standing of its members as their prestige raises the reputation and influence of the group.

Since family names pass on down the male line and children take the surname of their father [4], families from the same lineage share the same surname as the first ancestor. However, families with the same surname do not necessary belong to the same lineage since they might have a different pedigree. In our data set, since we can only observe the family name distribution, we must assume those with the same surname are from the same lineage. This is of little consequence as we focus on data from small villages, where a surname almost always implies a lineage. We therefore use family name and lineage as synonymous in this paper.

Most Chinese villages have hundreds of years of history, beginning when the first ancestor of a lineage founded the settlement. During that period, others moved to the village and began other lineages. Reliable historical statistics are, of course, nonexistent, but support can be drawn from the anthropology literature. Jing [1996] described a typical village in a rural county of Gangsu Province:

> Interpersonal relations in Dachuan (village) are tightly knit by descent and marriage. Although it is not a single-surname village, 85 percent of the local households in 1992 were surnamed Kong. The balance comprised 16 other surnames. Except for a group surnamed Li, whose ancestors had settled in Dachuan earlier than the Kongs, the others come to the village as refugees from war and famine.
>
> These Kongs trace their ancestry to Confucius through a Guangdong born ancestor who migrated to Gansu and settled in (this

[4] It only happens under extremely rare cases that a boy is surnamed with his mother's family name. Actually, the author never observed it in his own village, Liji Chun at Anyi County, with a population around 400.

> village) six centuries ago...the Kongs had long identified themselves with Confucian heritage by three means: a carefully guarded collection of genealogical records, a cycle of ancestral rituals and a temple dedicated to Confucius.

The above citation indicates that kinship relation still plays an important role in today's Chinese villages despite the fact that lineage in the traditional sense of ownership of common assets had not existed since the establishment of the People's Republic of China. Note that the first ancestor of Kong's lineage moved to the village almost six hundreds years ago after lineage Li's ancestor, and the ancestors of families with different surnames sequentially moved to the village because of war and famine.

With this evidence, we conclude that the lineage or family name distribution at the village level is determined historically. We next show that contemporary economic conditions do not influence the distribution. Without time series data on it, this cannot be shown effectively. Alternatively, the claimed exogeneity is built by showing that families from different lineages or income groups have the same amount of boys. The focus on the amount of boys is justified by the fact that only boys usually stay in the village when they grow up and thus are considered as part of a lineage. Furthermore, the strictly regulated residence permit *hukou* policy reinforces the claim since villagers cannot freely migrate between different villages.

One speculation that could potentially invalidate the exogeneity assumption is that richer families in a village could have more boys than other families. If this is true, the lineage distribution would depend on economic status. To test this hypothesis, we use the data set China Household Income Project(CHIP) 2002, which includes detailed information on family member and household income. The coefficients(standard error) of the regression of the number of boys on household/head income are 0.05(0.13)/0.06(0.13). These are not significant. The conventional wisdom that wealthier families can have more children or boys does not apply for two reasons. The one child policy makes it very expensive to have more children, especially more boys (Ebenstein [2008]). The prevalent, strong son preference in rural Chinese villages provides incentives for a poor family to have a boy by all means. The Confucian doctrine of filial piety teaches "there are three ways to be unfilial, the worst is not to produce male offspring". A traditional Chinese family would employ all of its assets to have a son. This is verified by the finding that wealthier families do have more children but not more boys.

The second conjecture is migration between different villages as a response to economic circumstances. However, this is not warranted since lineage is based on descent relationships and migration definitely cuts this relationships without much benefit (see more on inter lineage relationships in the next section). The strict residence regulation policy, *hukou*, since 1950s closes the door for any migrations on grounds of personal economic considerations (Chan and Zhang [1999]).

In summary, the historical and empirical evidence reasonably supports our claim that surname distribution is exogenous with regard to current economic conditions. The last subsections build the first stage relationship by presenting the inter-lineage rivalry and cooperation relationship, which are essential for constructing the instruments. During the discussion, the inter-lineage relations are limited to the lineages from the same village. A better viewpoint would be the inter-lineage relations among neighbour villages. The latter one is infeasible in our case because of the data limitation.

2.4.2 Inter-Lineage Rivalry and Cooperation

Understanding interactions between lineages is fundamental to understanding and capturing the effect of lineage composition on trusting behaviour at a village level. In this section, we briefly introduce some findings from anthropology about how lineages battle with each other and cooperate as a unit when it is necessary. The assumption is while serious rivalry impedes development of a trust, cooperation between different lineages builds it.

How rivalry affects trust has a long history of debate among social scientists. Intuitively, a fierce competitive market impedes reputation building [5] and deters trusting attitude. From a cultural evolution viewpoint, rivalry

[5] See Bowles [1998] for a comprehensive review.

can influence the structure of social interaction among lineages, therefore affecting social norms of parent-to-son trust transmission process(see Bisin and Verdier [2001]; Francois and Zabojnik [2001]). Though a positive relationship between trust level and competitiveness was reported in some research(*e.g.*, Francois and van Ypersele [2009]), rivalry differs from competition and does not foster a trusting attitude. Rivalry is more stressful and aggressive. And unlike competitors, rivals follow no rules.

Lineages vie not only for land or water, but sometimes prestige and *feng-shui*, which is believed to bring fortune, wealth, and health to the whole lineage. Beattie [1979] showed that rivalry among lineages in Tongcheng, Anhui province, for social prestige is a critical element for the increasing success of the lineages in the examination system. Baker [1979] noticed how the Confucianism doctrine of filial piety turns "family by being inward-turned ... outwardly aggressive." and "The same applied to the lineage: but the lineage was also expansionist." Baker [1968] also illustrated a lineage battle with other lineages for good *feng-shui* by tearing down a pagoda after incorrectly building it:

> It was thought that the open mouth of the eagle was swallowing some of the good fortune of the Liaos, so that male children and examination success in particular were being denied the lineage, and a geomancer suggested that between the village and the eagle a *feng-shui* pagoda should be built to represent a bird table, thus protecting the Liaos from the bird's appetite. Unfortunately the pagoda was not built on the correct site, and other villages to the west benefited instead, gaining unprecedented successes in examinations. The lineage tore down the pagoda...

The rivalry for land, water, and sometimes control over a local market even results in feuds. Freedman [1966] vividly described some feuds among lineages and concluded "...that fight between lineages... was an important characteristic of social life..." Hu [1983] gave a detailed description about how two lineages fought for a small mountain with woods, which finally claimed seven lives. Though loyalty encourages taking part in the fighting in a full scale feud, the rewards are provided by the lineage, which includes taking care of the families, widows, and immortalization ritual to those dead in the war.

However, some students (see Strauch [1983]; Johnson [1976]) observed that patrilineal ideology does exert an undeniable influence on inter-lineage relationships, but it does not mean a village compound of different lineages cannot work together in times of need. It seems reasonable to assume that lineage solidarity can coexist harmoniously with community solidarity from an infinitely repeated game perspective. The field work by Strauch [1983] in a multi-lineage village named *FungYuen* supports this point. In the village, a sense of village membership is at least as strong as the lineage identification and village exogamy is strictly enforced. Johnson concluded that:

> The people ...lived in an immediate social world populated chiefly by themselves, and they developed patterns of cooperation in economics and ritual spheres that united them even more strongly than their separate ancestral loyalties divided them.

There are many reasons for strong cooperation rather than rivalry in village life. Risk sharing or cost sharing are among the most prevalent. For example, lineages in the same village would unite to compete with other villages for a good tutor for children. To defeat robbers and pirates, families from different lineages come together. Strauch [1983] reported how people from different lineages formed groups and shared resources during the Japanese occupation period in the village. Baker [1966] detailed how lineages in Hong Kong draw together and cooperate after being attacked by pirates:

> ... but rather lineages of the five clans which came together and each purchased a share in the temple.... Not only was land purchased and a temple built with this money, but also a ferry boat was bought to assist all members of the five clans to cross the Sham Chun River to get to the large market town of Sham Chun, with which all had dealings. The share-holding lineages

took part in an annual feast at which the business of the temple was discussed, the feast being paid for out of temple funds.

From the above discussion, it seems safe to conclude that rivalry and cooperation among lineages are not mutually exclusive in village life. The coexistence of multifaceted social interaction strategies makes it unclear how the lineage composition affects social capital. Thus, it is necessary to use different methods to capture the effect of rivalry and cooperation on social capital. For this reason, we next discuss how to construct two instruments to respectively capture rivalry and cooperation among lineages.

2.4.3 Measurement

This section constructs instrumental variables from a village's lineage composition. The key idea is to capture the two sides of the inter-lineage relations by proposing two different measurements. Concretely, we employ the fractionalization or Herfindahl-Hirschman index to capture rivalry and the polarization index to capture cooperation. The polarization index, though designed to capture total social antagonism of a polarized society, measures mutual affection from cooperation quite well. Let us denote the lineage distribution in a village as $(\boldsymbol{\pi}, \mathbf{1}) = (\pi_{1,} 1_1; ...; \pi_n, 1_n)$, where the positive integer n is total lineages in a village, π_i is household or population proportion of the lineage i in the village, which satisfies $\sum_{i=1}^{n} \pi_i = 1$, and 1_i is an index for lineage i.

Before presenting the formal analysis, Table 2.3 illustrates the association between social capital and family name composition. Columns 1 and 2 show the OLS estimates of social capital on the share of the most common surname while columns 3 and 4 list the results on the aggregated share of surnames other than the top three. The relationship is rough, but the greater the prevalence of the most common surname, the more inclined toward higher social capital. In contrast, the greater of the percentage of surname other than the top three, the more likely toward lower social capital. Though the exact channel of the association is not yet clear, it reveals the basic link between the surname composition and social capital.

Table 2.3: Families Name Composition and Social Capital

	(1)	(2)	(3)	(4)
Family Name Percentage	.54(.17)***	.61(.17)***	-.004(.001)**	-.005(.001)**
R^2	.05	.10	.04	.08
Obs.	395	393	358	350
Village Charac.		YES		YES

Notes: The dependent variable is social capital as measured by trusting attitude. In columns 1 and 2, family name percentage is the the percentage of the most common surname, while in columns 3 and 4 is the share of the surnames other than the top three. The standard errors are in parentheses. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

The fractionalization index (FI), defined as $\sum_{i=1}^{n} \pi_i(1-\pi_i) = 1 - \sum_{i=1}^{n} \pi_{i^2}$, has been widely adopted in conflict research to examine the relationship between ethnic diversity and poor economic performance resulting from investment deterrence, impediments to technology innovation, or potential conflict (see Easterly and Levine [1997]; Alesina *et al.* [2003]). The index is commonly interpreted as the probability of two randomly selected individuals will not be of the same group. Alesina and Farrara [2002] reported a significant negative correlation between the ethnic fractionalization index and trust levels. Vigdor [2002] gave a simple behavioral model to provide an economic motivation for fractionalization effects.

To capture the cooperation effect, we use the polarization index (PI) developed by Esteban and Ray [1994] and Montalvo and Reynal-Querol [2002]. By assuming that a population of individuals may cluster according to some characteristics and clusters are mutually antagonistic, the polarization index was originally designed to measure polarization through total social antagonism. Since we are trying to measure how cooperation affects trust, instead of assuming antagonism amid different lineages, we assume that cooperation among lineages brings mutual affection. This is appropriate considering social capital is measured by the average trust level in strangers in our data set. Thus, our polarization index measures the total inter-lineage affection in a village.

The construction follows Esteban and Ray [2004], except we change antagonism to friendship. Assume that intra-lineage identification weakens inter-lineage friendship and inter-lineage cooperation accentuates affectionate attitude in the village. Specifically, let the identification function $I : [0, 1] \rightarrow R$ and cooperation function $c(d(1_i, 1_j))$ be an increasing function with properties $c(0) = 0$, where $d(1_i, 1_j)$ is the social distance between two lineages i and j. To capture the affection that any individual from lineage i feels towards lineage j, we define the effective friendship function $F(I, c)$ to be strictly increasing in c. The total polarization in a village level is defined as the sum of all the effective friendship: $P(\boldsymbol{\pi}, \mathbf{1}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i \pi_j F(I(\pi_i), c(d(1_i, 1_j)))$. Under some reasonable axioms, it can be shown that the polarization index can be expressed as $P(\boldsymbol{\pi}, \mathbf{1} : \tau) = K \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i^{1+\tau} \pi_j d(1_i, 1_j)$, where $K > 0$ and $\tau \in (0, 1.6]$ is the degree of polarization sensitivity. Following Montalvo and Reynal-Querol [2005], we further substitute the social distance $d(1_i, 1_j) = 1$ if $i \neq j$ and $d(1_i, 1_j) = 0$ otherwise, and the sensitivity parameter $\tau = 1$. This assumption still allows for intra-lineage friendship. It is maintained because we want to measure the effect of inter-lineage cooperation on social capital. Finally, our polarization index is $PI = 4 * \sum_{i=1}^{n} \pi_i^2 (1 - \pi_i) = 1 - \sum_{i=1}^{n} (\frac{0.5 - \pi_i}{.05})^2 \pi_i$ [6].

The CGSS data set only includes the top three lineages' proportion for each village, thus some of the distribution is unobservable, truncated beyond the top three lineages. The total share of the top three lineages is less than 85 percent for two thirds of the villages in the data set. Ignoring the data truncation can introduce serious bias. We take two steps to deal with this problem. First, for the villages truncated beyond the third biggest lineage, we assign the residual portion to other lineages equally; the number of unobserved lineages is calculated based on the shares of the third largest lineage. For example, consider a village with the lineage composition $(\boldsymbol{\pi}) = (\pi_1 = 40\%, \pi_2 = 20\%, \pi_3 = 15\%)$. The sum of the top three lineage is $40 + 20 + 15 = 75$ percent. The residual share is $100 - 75 = 25$ percent. The

[6] Actually, $\tau = 1$ is the only level generating the polarization index that satisfies some good properties of polarization and $K = 4$ is the only value such that the index ranges from zero to one. See Montalvo and Reynal-Querol [2002] for detailed proofs.

number of the non-observed lineages is calculated by the ceiling integral of $25/15+1$ which is 3. So, the residual share is equally divided to the fourth, fifth and sixth lineages and each lineage has a share of $25/3 = 8.34$. The second step is introduced in section 2.5.

This method exploits the rank information included in the data set. It relies on the assumptions about the unobserved lineages. However, compared with other methods that rely on distributional assumptions, we believe it is more acceptable since it makes the least assumptions. Different methods that assume unequal distribution among the left proportion were tried as well, but no differences were found and thus these results are not reported here.

Figure 2.2 presents the relationship between the FI capturing rivalry and PI capturing cooperation. Obviously, FI has a horizontal U-shaped relationship with PI. In other words, PI maximizes when FI is around 0.5, and minimizes when FI is around 0 or 1. A FI around 0 or 1 implies a village consisting of either one dominating lineage or many small lineages. Under either situation, rivalry force will be taken over by a need for cooperation. Similarly, a FI around .05 must be from villages with several dominating lineages, a situation freezing cooperation.

2.4.4 First Stage Analysis

The PI index captures cooperation and FI index measures rivalry. Thus, social capital as measured as trusting attitude should have a positive correlation with PI and a negative one with FI.

Figure 2.3 plots the partial correlation between social capital and the constructed instruments at the village level, where the upper panel is for the rivalry index and the lower for the cooperation index. Notice a similar figure at the individual level should have the same pattern except a tighter confidence interval for the fitted line and thus is omitted. The visual representation shows a strong first stage relationship between social capital and instruments: a negative partial correlation for the FI and a positive one for the PI as expected. The higher the FI, the more lineages concentrated,

Figure 2.2: Fractionalization *vs.* Polarization Index

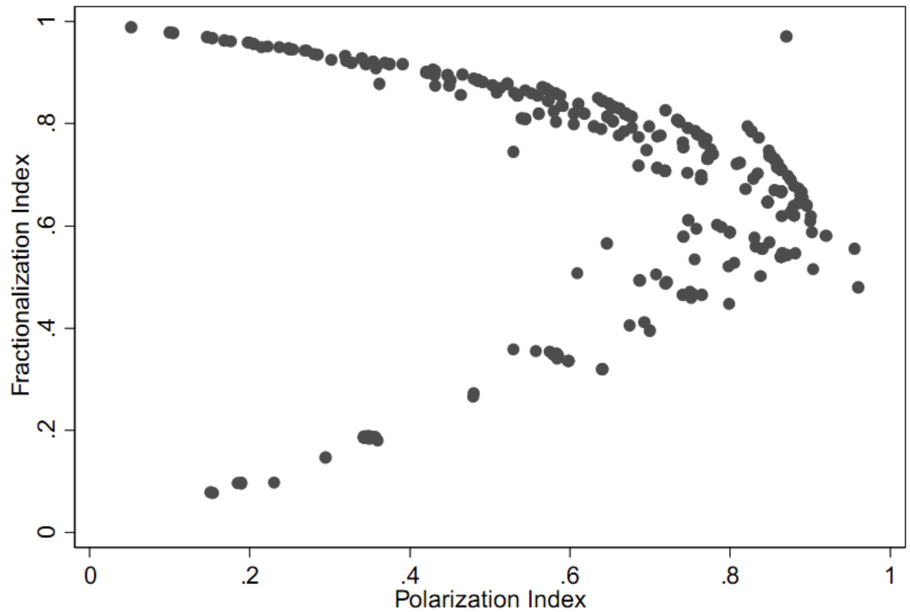
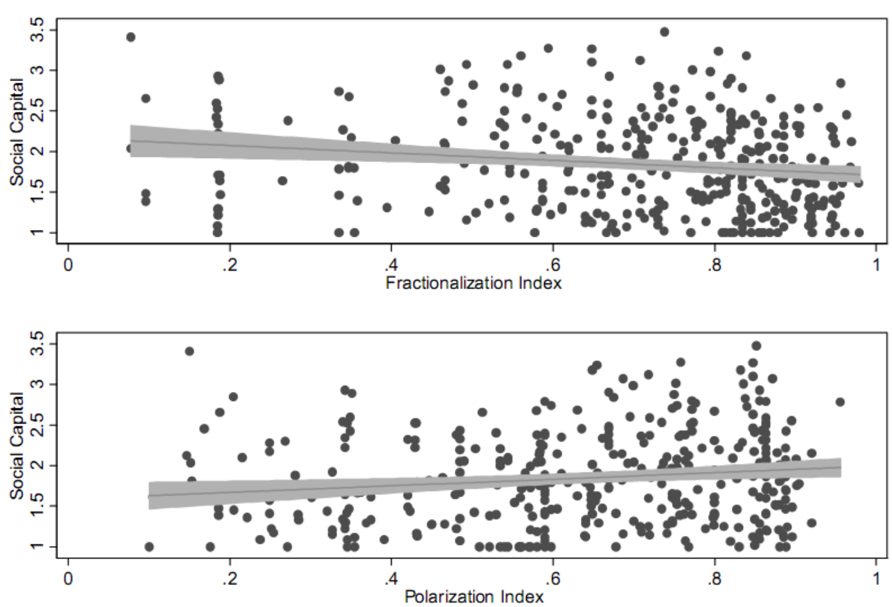


Figure 2.3: Partial Correlation of Social Capital and Instruments



lead to more rivalry. Cooperation is reinforced in a more "polarized" village, as suggested by a higher value of PI. The significant correlation adds confidence in the validity of the instruments to capture rivalry and cooperation among lineages.

Table 4.1 details the regressions from the first stage at village level. Panel A pertains to the FI and panel B to the PI. Two features deserve special attention. First, the regression verifies the claim that the constructed instruments have a significant effect on social capital subject to controlling various covariates. Technically speaking, what is relevant to the analysis is the individual level first stages. The village level can be treated as a robustness check since it is possible the strong individual correlation between the instruments and social capital might mainly result from repeated observations. The estimates in the table eliminate such concern. Individual level results are available and support the above analysis. Second, a striking feature in the table is that $F-$values in the first stage regressions are lower than what is needed for a strong instrument, say an $F-$value of 10 suggested by Stock, Wright and Yogo [2005]. This is true even in the individual level regression. Thus, the weak instrument problem is a concern and should be treated carefully.

Overall, Figure 2.3 and Table 4.1 show that there are significant first stages for social capital at the village level, but the first stages potentially suffer from the weak instrument problem.

2.4.5 First Stage Robustness

The first stages show our constructed instruments have a significant relationship with social capital as expected. Nevertheless, our measurement of social capital is based on the decisions made by subjects (to what extent to trust strangers). It is possible what measured is not social capital, but something unobservable in the data that affects the decisions. To address this concern, we perform two robustness tests based on the origins of lineage and anthropological evidence.

The connection between family name and lineage is mainly based on

Table 2.4: First Stage Regression for Social Capital

	(1)	(2)	(3)
	Panel A: Fractionalization Index		
Fractionalization Index	-.50(.19)**	-.52(.19)**	-.57(.19)***
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.034	.040	.013
$F-$value	6.75	1.73	3.60
Observations	378	375	344
	Panel B: Polarization Index		
Polarization Index	.28(.17)*	.28(.17)*	.31(.17)*
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.010	.013	.099
$F-$value	2.83	.81	3.03
Observations	378	375	344

Notes: Standard errors are in parentheses. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is social capital measured by trust in strangers at village level.

Confucian philosophy, so is the first stage relation. It is standard to assume Confucianism provides the core values and norms of Chinese society, but these values are actually practiced mostly by Han Chinese. It is questionable to what extent Confucian values can be generalized to other ethic groups. Shein [1994] illustrated how a member of the Miao Chinese is so different from the Han Chinese in terms of culture, including ritual activities and spiritual beliefs *etc.*. Thus, our first test is to divide villages into two types; those that only include Han Chinese and those populated with all other Chinese. To simply the analysis, any village with at least one minority peasant is treated as a minority village. In total, 59 villages out of 318 are categorized as minority villages.

Secondly, anthropologists like Freedman [1966] have argued the agricultural surplus associated with a rice economy initially made possible the establishment of corporate estates which in turn promoted the development of large patrilineal communities. He noted a correspondence between regions with large, localized lineages and areas of rice cultivation. Pasternak [1969] extended this claim and argued that the exigencies of frontier life were crucial for the emergence of strong lineages. There is, in addition to the cooperative effort necessary to bring wild land under cultivation, a need for organized defense against bandits invaders. Rice cultivation and frontier life in Chinese history were generally a phenomena of South China. From this evidence, it seems reasonable to conclude that lineages are more stable and effective in south China. Our second test differentiates villages form South and non-South, including North and West China [7].

Table 2.5 and Table 2.6 present the estimates. The dependent variable *Fractionalization Index*Minority* is the interaction item of Fractionaliztion Index and Minority dummy variable. Similarly for other interaction items. Both panel A and panel B show that in the minority villages where the lineage system is very weak or nonexistent, the interactions have no essential effect on social capital. They also give a significant and robust first stage

[7] By South China, we refer to these provinces: Jiangsu, Zhejiang, Anhui, Jiangxi, Shangdong, Fujian, Guangdong, Guangxi, Hubei and Hunan. All other provinces in our data are treated as non-South China.

Table 2.5: First Stage Estimates: Minority Villages *vs.* Han Villages

	(1)	(2)	(3)
	Panel A: Fractionalization Index		
Fractionalization Index*Minority	-.10(.45)	.-.08(.40)	-.15(.70)
Fractionalization Index	-.51(.20)**	-.54(.22)**	-.47(.17)**
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.25	.28	.40
Observations	378	354	344
	Panel B: Polarization Index		
Polarization Index*Minority	.07(.44)	.15(.38)	-.10(.47)
Polarization Index	.32(.17)*	.42(.17)**	.40(.16)**
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.20	.23	.34
Observations	378	354	344

Notes: Standard errors are in parentheses. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is social capital measured by trust in strangers at village level.

relationship for those villages consisting of both Han and minority Chinese where the lineage system averagely has a tradition and is rooted into daily life and spiritual beliefs [8].

Table 2.6 has a similar structure. The interaction items are the constructed instrumental variables and the South or North dummy variable. The similar pattern of the estimates brings us more confidence in the instruments. In both cases, our results negate the possibility that the first stages presented in the last section are driven by other unobservables.

These results therefore indicate that the relationship in the first stage are not likely driven by unobservable variables, which gives us more confidence in using the constructed instruments to investigate the effect of social capital on individual economic achievement.

[8] This conclusion is even more apparent when we divide the sample into North and South (Han vs. Minority) groups and compare the estimates from these groups.

Table 2.6: First Stage Estimates: South *vs.* Non-south

	(1)	(2)	(3)
	Panel A: Fractionalization Index		
Fractionalization Index *North	.04(.27)	-.03(.26)	-.10(.18)
Fractionalization Index	-.66(.23)***	-.65(.25)**	-.71(.23)***
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.23	.25	.29
Observations	378	378	344
	Panel B: Polarization Index		
Polarization Index*North	-.07(.28)	-.04(.36)	-.03(.25)
Polarization Index	.44(.21)*	.43(.18)**	.39(.20)*
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.20	.25	.27
Observations	378	378	344

Notes: Standard errors are in parentheses. A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is social capital measured by trust in strangers at village level.

First Stage Mechanism So far, our first stage analysis has shown evidence for the strong relationship between social capital and inter-lineages relationship. Since social capital is measured as a trust level in strangers, the strong relationship demands an explore why the inter-lineages relationship that most likely happens among acquaintance could affect villagers' trusting attitude towards strangers.

Inter-lineage relationship, where villagers and their offspring can be modeled as playing an infinitely repeated game, affects players' prior that an opponent is cooperative rather than fully rational. In an experimental environment to examine how exogenously determined past relationship lengths affect trust and trustworthiness, Engle-Warnick and Slonim [2003] found that there is significantly more trust after long relationships. This is especially relevant in the Chinese rural villages where inter-lineage rivalry and cooperation dominated the rural social life in history.

The diffuseness of trust in neighbours to outsiders is more likely a result of multi factors, though reputation can be of more effective. Manapata, Nowaka and Randa [2012] discussed in the traditional trust game [9] how a little information on the trustee's reputation can lead to robust evolution of trust and trustworthiness through partner choice. In our model, with the acceptable assumption that a stranger (investors) have some knowledge of a trustworthiness of a villager (trustee), the result can be applied to explain the diffuseness of trust.

To test the general idea of free market exchange that might boost trust in strangers and general public, we decompose villager's income into two parts: one from agriculture production and the other one from non-agricultural market activity (including small business and non-agricultural labour income). Define market intensity as the ration of non-agricultural income to agricultural income and we find the correlation between market intensity and individually reported trust in strangers is around 0.43. This result seems

[9] A trust game consists of an investor and a trustee. The investor, endowed with some money, makes the first move by either keeping the money or transferring it to the trustee. The value created by interactions based on trust is the money multiplied by a factor $\alpha > 1$. Finally, the trustee chooses how much to return to the investor and how much to be retained.

support that market exchange might help to diffuse inward trust toward outsiders.

2.5 IV Estimates

2.5.1 Main Results

We previously mentioned that we need a second step to compensate for the truncated observation problem. The first step is used to impute the missing values which are essential to construct instruments. Its disadvantage is that the procedure is vulnerable to interference from noise, as we observed that among two thirds of villages the top three lineages are less than 85% of the population. The noise prevents the instruments from possessing significant variation in the instrumented social capital, making it hard to capture the effect of interest. To overcome this problem, we generate eleven dummies by the total proportion of the top three surnames and multiply these dummies by FI (PI). The first dummy is valued 1 if the total share of the top three surnames is from 1 percent to 9 percent. Others are defined similarly until the last one which is valued 1 if the total share is 100 percent. For example, consider the dummy tenth which equals 1 if a village's total percentage of the top three ranges from 90 percent to 99 percent and 0 otherwise. We generate the dummies according to the sum of the top three because the noise comes from the imputed values of the unobservable surnames. This hinges on the sum of the top three and the value of the third lineage. Since instruments are more than one, to get efficient estimates, we adopt two step GMM estimation. It is more efficient than two stage least squares since the former can capture heteroscedasticity.

Table 2.7 reports the results for males and females respectively. Panel A uses the FI as the instrument while panel B uses the PI. Column (1) does not include any controls. Column (2) only includes controls for individual characteristics like age, square of age, education and parents' education. Column (3) additionally includes average individual characteristics such as village average education and village average age and its square. Column

(4) includes other village characteristics listed in table 2.1. In column (1) of both panels, the coefficients for males are around 0.50 and statistically significant at 10% level. The coefficients for females are neither stable nor significant. Column (2) to column (4) illustrates this first order effect by controlling various individual and village characteristics, the estimates range from 0.34 to 0.56. The standard deviation of social capital is about 0.57 in the data set, which means an increase in one standard deviation in social capital can increase the villagers' incomes by 19% to 32%. Column (5) lists the estimates with the average village characteristics that do not include the own-person characteristics. Generally, the return of education is about 8% to 10%. So, a one standard deviation increase in social capital was found to be equivalent to about two to four years of education. As we found in the OLS estimates, the IV estimates also indicate the absence of any effect on female villagers.

Table 2.7 also reports F values from the first stage regression. The F values are lower than the conventional recommended value of 10 for a non weak instrument. This implies a very low value of the concentration parameter, that is the proportion of the social capital's variation explained by the included variables. Actually, we have encountered this problem in the first stage analysis earlier. As discussed by Stock, Wright and Yogo [2002] and Stock and Yogo [2005] among many others, the estimates can be biased and the standard inferences based on normal distribution can be misleading. Thus, caution should be applied in reading the results in the table [10].

Overall, the results in this subsection suggest that social capital has a first order effect on income for males, but not for females. Holding all other factors constant, a male villager can increase his income more than 19% to 32% if he is from a village with one standard deviation higher social capital, which is averagely equivalent to 3 years of education. This result is robust in the sense that it holds under various specifications. However, the first stage F values also imply a potential weak instruments problem, which could contaminates the conclusion.

[10] Another potential issue is the small sample bias for IV. As Ham, Kagel and Lehrer [2005] shows, a sample size of 3000 is necessary for a good inference.

Table 2.7: Effect of Social Capital on Income: IV Estimates

	(1)	(2)	(3)	(4)	(5)
Panel A: Instrument Variables: Fractionalization Indices					
	Male				
Social Capital	.50(.28)*	.56(.25)**	.57(.25)**	.35(.17)**	.37(.16)**
R^2	.01	.03	.04	.14	.15
$F-Value$	2.10	2.00	1.72	2.02	2.34
Obs.	1247	1156	1134	1113	1113
	Female				
Social Capital	-.08(.18)	.03(.16)	.03(.16)	.01(.17)	0.02(.13)
R^2	.01	.05	.09	.10	.12
$F-Value$	1.25	1.19	1.49	1.61	1.73
Obs.	1125	1034	1031	1001	1001
Panel B: Instrument Variables: Polarization Indices					
	Male				
Social Capital	.49(.26)*	.49(.23)**	.50(.22)**	.36(.18)**	.45(.19)**
R^2	. 01	.06	.07	.14	.16
$F-Value$	2.17	1.94	1.68	1.94	2.31
Obs.	1247	1156	1134	1113	1113
	Female				
Social Capital	.23(.25)	.20(.16)	.16(.17)	.13(.17)	.15(.23)
R^2	.01	.02	.08	.09	0.09
$F-Value$	.99	1.12	1.51	1.53	1.57
Obs.	1125	1034	1031	1001	1001
Controls:					
Individual Charac.		YES	YES	YES	YES
Average Charac.			YES	YES	YES
Village Charac.				YES	YES

Notes: Standard errors are in parentheses A single asterix denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is individual income. F-Value is the F-value at the first stage regression.

2.5.2 Robustness to Weak IV

In the literature, many methods have been proposed to correct the bias of weak instruments. However, most of these methods focus on the case of *i.i.d.* errors. In this paper, we apply a simple robust inference procedure following Chernozhukov and Hansen [2008]. Our main concern is whether the coefficient on social capital is some given values. Under the null that the coefficient is equal to a given value, the exclusion restriction implies that the coefficients on the excluded instruments in the reduced form should equal zero. Thus, testing the hypothesis that the coefficients of all the excluded instruments are zero is equivalent to a test of whether the coefficient of social capital equals the given value [11]. Chernozhukov and Hansen [2005] showed that the conventional Wald statistics for testing this hypothesis is asymptotically distributed as a χ^2 regardless of the strength of the excluded instruments. To find the weak IV robust interval, we repeat the above procedure for different values of the coefficient of the social capital. Concretely, we consider the coefficient intervals of $[-1, 1]$ with step 0.001. The merit of this procedure is to avoid the *i.i.d.* assumption required by many other methods. Notice we have more instruments than the endogenous variables. The constructed Wald statistics used to test the hypothesis also tests the specification.

Table 2.8 presents the weak IV robust intervals for the coefficient of social capital at the 95% confidence level. As a comparison, we also replicate the estimates in Table 2.7 and its 95% confidence interval. Panel A gives the results using the FI as the instruments. For males, the weak IV robust interval at the 95% confidence level ranges from 0.025 to 0.603. The estimated intervals robust to weak IV are generally tighter than the interval constructed the conventional way. Also, the two step estimates are close to the maximum value of the weak robust interval. For female villagers,

[11] To make it clear, consider the following structural model: $Y = SC\beta_1 + X\beta_2 + \varepsilon$ and $SC = Z\Pi + V$, where Y is the income, SC represents social capital, Z is the excluded instruments and X is the included instruments. The test of $\beta_1 = \beta_1^0$ is equivalent to the test $\gamma = 0$ under the reduced form $Y - SC\beta_1^0 = Z\gamma + X\beta_2 + \psi$. Notice here no information about the relationship between SC and Z is ever used.

Table 2.8: Effect of Social Capital on Income: Robust to Weak IV

	(1)	(2)	(3)	(4)
Panel A: Instrument Variables: Fractionalization Indices				
	Male			
Social Capital	.50(.28)*	.56(.25)**	.57(.25)**	.35(.17)**
Asymptotic Interval	(-.049,1.054)	(.045,1.071)	(.083,1.054)	(.003,.684)
Weak IV Robust Interval	(.025,.506)	(.235,.603)	(.254,.555)	(.129,.498)
Obs.	1247	1156	1134	1113
	Female			
Social Capital	-.08(.18)	.03(.16)	.03(.16)	.01(.17)
Asymptotic Interval	(-.437,.273)	(-.278,.336)	(-.288,.343)	(-.323,.332)
Weak IV Robust Interval	(-.999,1.001)	(-.999,1.001)	(-.999,1.001)	(-.999,1.001)
Obs.	1125	1034	1031	1001
Panel B: Instrument Variables: Polarization Indices				
	Male			
Social Capital	.49(.26)*	.49(.23)**	.50(.22)**	.36(.18)**
Asymptotic Interval	(-.017,.1.007)	(.053,.943)	(.095,.941)	(.016,.676)
Weak IV Robust Interval	(.210,.401)	(.283,.644)	(.274,.549)	(.095,.472)
Obs.	1247	1156	1134	1113
	Female			
Social Capital	.23(.25)	.20(.16)	.16(.16)	.13(.17)
Asymptotic Interval	(-.208,.455)	(-.112,.531)	(-.158,.487)	(-.205,.475)
Weak IV Robust Interval	(-.999,1.001)	(-.999,1.001)	(-.999,1.001)	(-.999,1.001)
Obs.	1125	1034	1031	1001
Controls:				
Individual Charac.		YES	YES	YES
Average Charac.			YES	YES
Village Charac.				YES

Note: This table replicates the Table 2.7. The asymptotic interval is the 95% confidence interval using the usual asymptotic approximation. The weak IV robustness interval reports the 95% confidence interval using the weak-instrument robust statistics.

the weak robust interval ranges form -0.999 to 1.001, the whole considered interval. These results establish the gender differentiation in social capital. Panel B displays the results where PI are used as instruments. The weak robust interval for male villagers ranges from 0.095 to 0.644, very close the interval found earlier. The female results are also unchanged.

Thus, the weak IV robust intervals verify our finding from two step IV estimates. It reinforces the claim that social capital has a significant and robust first order effect on males, but no detectable effect on females. The effect on male villagers varies from 2.5% to 64%, depending on the specifications. The interval becomes much tighter, from 9.5% to 49.5%, in column (4) where all related factors are controlled. Under the assumption that social capital has no direct effects on these factors, this interval implies that a one standard deviation unit increase in social capital is equivalent to one to three years' schooling.

2.5.3 Robustness to Too Much Noise

At the beginning of this section, we discussed why and how to generate eleven indices to capture variation in social capital. The general IV estimates and the estimates robust to weak IV show a significant first order effect. However, another concern is that the estimates we got in last the subsections are mainly driven by the eleven dummies, not by the instruments or alternatively, the estimates are mostly driven from the use of too many instruments. The implicit argument for the latter concern is that given so many instruments with so much noise, it is possible to get an illusionary first order effect.

The first argument can be tested by using only the eleven dummies in the IV estimation. The regression results show that without multiplying with the original instruments, the coefficients for both male and female are insignificant under various sets of controls.

To assess the second argument, we use the following randomization test. We first collect all the villages' family name distributions and then randomly match a village with one of these distributions. With this randomly gener-

Table 2.9: Effect of Social Capital on Income: Randomization Test

	(1)	(2)	(3)	(4)
Panel A: Instrument Variables: Fractionalization Indices				
	Male			
Social Capital	008(.445)	.007(.476)	.002(.467)	.014(.356)
(min, max)	(-2.302,1.350)	(-1.555,1.282)	(-1.345,1.456)	(-1.387,.989)
Panel B: Instrument Variables: Polarization Indices				
	Male			
Social Capital	-.016(.440)	.011(.469)	.035(.454)	-.007(.342)
(min, max)	(-1.190,1.720)	(-1.403,1.393)	(-1.879,1.385)	(-1.168,1.351)
Controls:				
Individual Charac.		YES	YES	YES
Average Charac.			YES	YES
Village Charac.				YES

Notes: This table repeats the randomization described in section 2.5.3 one thousand times. The coefficient and thus the standard errors calculated as the mean and standard deviation of the one thousand estimates. The dependent variable is income.

ated data set, we perform the IV estimation as we did at the subsection 2.5. We repeated this one thousand times and recorded the coefficient from each randomization. Finally, we use the mean and standard deviation of these coefficients as the coefficient and its standard error of the randomization test. Note this method is very similar to the classical bootstrap estimation, but differs in how to generate the repeated observations. The randomization test has a good size in the sense that if the second argument is true, the randomized coefficients should be significant as well.

The results for male are listed in Table 2.9. It clearly indicates that we cannot reject the assumption that the true value of the coefficients are zero under various controls. Thus, Table 2.9 negates the argument that our results are possibly driven by the noise in our data set.

2.5.4 The Exclusion Assumption

The validity of the identification strategy that was used in the above section rests on the assumption that family name composition is a legitimate instrument for social capital in the earnings equation. It becomes clear in the first stages that the family name composition is related to social capital because of lineage. However, for the validity of the IV strategy, it must be true that the family composition is uncorrelated with the residual in the income equation 2.2. In other words, if family name composition affects income other than social capital, our approach is called into question. We argue the exclusion assumption holds for the following reasons:

First, as it has been shown in the first stage robustness analysis where Han *vs.* non-Han and south *vs.* non-south Chinese villages were employed, lineage does affect social capital, even though as an institution lineage was abolished by the Communist party in 1950s and replaced by village party branches.

Second, it is necessary to distinguish between family names and family name distribution. It would be true that the family name of a villager can affect her personal earning in some way other than through social capital, but it does not necessary indicate that the distribution of family names is significantly correlated with income through other channels.

Third, it is possible that family name composition could affect earnings through effects on the provision of public goods and risk sharing contracts *etc.*. If family name distribution is correlated with a village's capacity to provide public goods or sign risk sharing contracts, the exclusion assumption is not warranted. However, its effects on estimates are not clear due to the fact that the instrumental variables in our analysis are constructed from the distribution of family names. Further, if such effects are substantial, we should expect substantial differentials among the estimates from fractionalization index and polarization index since these index are nonlinearly correlated. However, the IV estimates clearly shows that is not the case.

Though risk sharing is difficult to measure, the CGSS 2005 data set does provide a good measurement on village level public goods provision, which includes expenditure on roads and bridges, agriculture infrastructure, electricity and communication *etc.*. To test this possibility, Table 2.10 reports the OLS estimates of public expenditure on the constructed instruments at village level. None of the coefficients of either instruments is significant.

Table 2.10: Does Family Name Distribution Affects Public Goods Provision?

	(1)	(2)	(3)
	Panel A: Fractionalization Index		
Fractionalization Index	.34(.57)	.14(.59)	-.08(.63)
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.001	.033	.013
$F-$value	0.3	1.32	2.97
Observations	378	375	344
	Panel B: Polarization Index		
Polarization Index	-.41(.63)	-.19(.61)	-.04(.63)
Controls:			
Average Charac.		YES	YES
Village Charac.			YES
R^2	.002	.04	.13
$F-$value	.43	1.25	3.05
Observations	378	375	344

Notes: Standard errors are in parentheses. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%. The dependent variable is public good expenditure at village level.

Also notice the low F values at all regressions. Though this result can not completely verify the exclusion assumption, it does show that, at least from public good provision perspective, the effect would be very small.

2.5.5 Interpretation of the Gender Difference

Our empirical investigation reveals a striking result: social capital has a large first order effect on male villagers but no real effect on female villagers. Though a full explanation is difficult, some possible causes can be proposed from both the lineage practice and the motivating model.

First, women might be less active in village affairs and therefore are less exposed to the lineage system because they usually are not born in their husbands' villages. Indeed, they migrate to the village for the purpose of marriage. Thus the constructed instruments can not capture the variation of social capital accessible to women since the validity of instruments relies on the claimed connection between family name composition and social capital. In other words, women are less influenced by the lineage system and naturally their income are less dependent on social capital.

Second, the motivating model does predict a significant effect of social capital, at least for male villagers. Thus, one way to understand the gender differentiation is to trace back to the model and speculate at which step we could change the model to produce the result observed. In other words, we suspect that the gender difference can be attributed to occupational differences, that is, male villagers mainly engage in business that requires exchange with a third party, *e.g.*, making a bed as a carpenter in exchange for money, while female villagers work in home production which does not involve a third party, *i.e.*, raising backyard chickens and selling eggs. Eggs-selling, which can also be done by husbands, is of course an exchange, however raising chickens, which usually is limited to wives, is the more essential part of the chicken business. Formally, male villagers' work is more social and thus reputation based and female villagers' work less social and thus less influenced by social capital. The above interpretation is merely a conjecture and more detailed work is necessary. Since the present data does not include occupational information, we leave this task to future research.

2.6 Conclusion

There is now considerable evidence that social capital, defined as trust in strangers or membership of associations, is an important determinant of economic performance and financial developments. For example, Guiso *et al.* [2004] identified the effect of social capital on financial development by exploiting differences in social capital within Italy. They found in high trust areas, households are more likely to use cheques, have higher access to institutional credit and rely less on informal credit, even after controlling for social environmental variables. As other work that links social capital and economic outcomes, research work investigating the causal effect of trust on individual economic achievement usually suffers from the endogeneity problem.

Our identification strategy is to exploit the variation in the family name distribution of Chinese rural villages. The unique cultural tradition of the lineages offers credible exogenous variation in social capital as measured as trust in strangers at village level. The abolishment of the lineage institution as a formal institution in the 1950s and our focus on the distribution of surnames reinforce the confidence in the exclusion assumption. We construct two indices from the distribution to instrument for village level social capital. The first stage analysis displays a significant and robust relationship between the indices and trust. Our GMM estimates mimic the OLS results and find that a one standard deviation change in trust is almost equivalent to two to four years of schooling in terms of the causal effect on earnings. However, this effect only exists among male villagers, not among female villagers. Robustness to weak IV estimates confirm the above conclusion.

Our conjecture for the gender differentiation in the economic effect of trust is that male villagers' earning are mostly from free market trade while female villagers' are mostly from home production. We build a simple model to show that trust determines a villager's trade volume and thus income. Since home production doesn't rely on trustworthiness, our model provides a simple explanation for the difference.

This paper is a first step to understanding trust and individual earnings. Needless to say, much more empirical and theoretical work is still required. For example, other historical events that might have some impact on social capital and income are not accounted in our analysis. These events include civil war, the Great Famine and the cultural revolution *etc.*. All of these can bias the results if they are correlated with the name distribution. No discussion to explore how the historical contest of lineage rivalry and cooperation can effect the current trust level. We adopted two indices, namely fractionalization and polarization, to capture the effects of lineage relations on trust. Are these two indices the most appropriate? If not, how to construct other indices to build reasonable behavior assumptions? We believe these topics are fruitful areas for future research.

Chapter 3

Identifying Subjective Expectation Bias: Method and Evidence from the Health and Retirement Study

3.1 Introduction

Economic models of choices under uncertainty typically assume people choose an action to maximize expected utility by combining their preferences and subjective probability distributions over uncertain outcomes. Without data rich enough to include subjective expectations, researchers need to further assume rational expectations in order to identify interesting preference parameters and predict choices. This compromise is necessary since the observed choice behaviour would be consistent with many alternative combinations of preference and expectation, which was elegantly discussed by Manski [2004]. Between theory and practice, one critical question is whether people form subjective expectations in the way described by the rational expectations assumption.

Empirical tests have generated mixed results. In an experimental environment, many psychologists have reported various subjective expectation biases. Tversky and Kahneman [1974] found various biases of judgment under uncertainty and Tversky and Kahneman [1981] reported that the evaluation of probabilities depends on the way a problem is framed. Hogarth [1975] offered a comprehensive review on cognitive processes and subjective probability assessments and concluded that people are ill-suited for assessing probability distributions because of limited capacity to process information. Though these researchers reported various abnormal observations on subjective probability evaluations, it is not clear how these abnormalities depart from rational expectations. In contrast, many economic researchers have found support for the rational expectations assumption. Bernheim [1988] found that seniors from the Retirement and Health Study responded rationally to new information, though the subjects did not extract all available information to form expectations. Smith, Taylor and Sloan [2001] evaluated the relationship between subjective beliefs on mortality and actual death at the individual level and found remarkable consistency. Benítez-Silva and Dwyer [2005] tested the retirement expectations of old married American couples and concluded they are consistent with the rational expectations hypothesis.

This paper proposes a simple definition of expectation bias by exploiting an implication of the rational expectations assumption. The rational expectations assumption states that people should hold an objectively correct expectation conditional on what information they possess (Manski [2004]). This implies that the same information should be weighted identically in projecting both objective and subjective probabilities. Accordingly, expectation bias is defined as the inequality of the weights.

To illustrate, consider the problem of seniors in the Health and Retirement Study (HRS) data set. The seniors were asked to assess their individual probability of entering a nursing home within five years. Consider the information represented by a variable reporting one's health status. To make things simple, let us assume health is measured properly. Under the rational expectations assumption, the coefficient of health status should be α in projecting the subjective probability, if the coefficient in projecting objective probability were α.

Of course, the objective probability could never be known outside a controlled experimental environment. What is observed is the realization of a random event. Thus, to assess rationality and identify expectation bias, we need to substitute the observed realizations for objective probabilities. The substitution brings an identification issue, which is whether the coefficients are still comparable between the equations of objective probability and of *ex post* realization, by introducing interim events between the formation of expectations and the realization of the random event. This paper shows that as long as an orthogonality condition holds, identification can be successfully achieved.

Although subjective probability can be elicited through survey questions, the measurement error could be a serious issue for any formal analysis. We find that the self-reported subjective probability of the seniors in the HRS dataset is measured reasonably well, which is consistent with many previous evaluations of the same dataset. For example, Hurd and McGarry [1995, 2002] found the HRS respondents' self-reported subjective probability of survival aggregates to actual population probability and predicts actual survival.

We find evidence for expectation biases among both female and male seniors. The female seniors do not correctly internalize age information into their subjective assessment of the probability, while the male seniors internalize income information poorly. The estimated coefficient of age [1] in projecting subjective probability for the female sample at various survey years is about 0.02 while the estimated coefficient in projecting objective probability is about 0.05. The estimated coefficient of income for the male sample in projecting subjective probability is about 0.01 while the corresponding coefficient in projecting objective probability is about $-.25$. The hypothesis of identical age coefficients is rejected at less than a 0.1% level of p-value and a similar hypothesis about the income coefficients is rejected at less than a 5% level of p-value.

Some potential threats to our conclusions are analyzed, which provides primary evidence against the suspicion that the identified biases could be driven by wishful thinking of seniors or a violation of the orthogonality assumption due to public policies regulating nursing home facilities. Discus-

[1] Which is measured as the true age in years minus 65.

sion on the sources of the biases shows that the age bias and income bias can only slightly be attributed to cognitive factors or a misconception between probability and risk aversion.

The remainder of the chapter is structured as follows. We propose the definition of expectation bias in section 3.2. To do so, we briefly review some related work before discussing the identification assumption. The estimation framework is also briefly presented. Section 3.3 provides a background to the data and a descriptive analysis. Choosing covariates and evaluating subjective probability data are also presented there. Section 3.4 gives the main results of the paper, which is followed by robustness analyses. Section 3.5 discusses two potential sources of identified biases. Section 3.6 concludes.

3.2 Definition and Identification Assumption

This section is the major theoretical part of the paper. It first briefly reviews related literature before formally proposing a definition of expectation bias. The definition compares the estimates of the *ex ante* objective and subjective distribution functions. Since the *ex ante* objective probability is not observable in most cases, our identification of bias hinges on a weak orthogonality assumption, which essentially states that the unexpected part of interim events between *ex ante* and *ex post* are orthogonal to population characteristics. Following that is a short discussion in model specification and equality hypothesis test.

3.2.1 A Belief Literature Review

Understanding how people form expectations is both interesting to students and important to policy makers. In academia, de Finetti [1969] stated "the true problem consists in the investigations concerning the ways in which probabilities are assessed by more or less educated people..." Since forward-looking individuals make decisions based on an assessment of unknown future states, better understanding its underlying mechanism can improve the efficiency of public policies in the fields of schooling (Manski and Dominitz

[1996] and Attanasio and Kaufmann [2009]), saving (Hamermesh [1985] and Dominitz and Manski [1997]) and retirement (Chan and Stevens [2004]) *etc.*. As Manski [2004] argued, it is theoretically impossible and practically inferior to infer preferences and predict future choices without invoking subjective probability. In experiments, Nyarko and Schotter [2002] and Bellemare, Kröger and Van Soest [2008] found models using subjective probability data can generate better sample predictions than various models where agents are presumed to form expectations rationally.

The work from different fields has produced different conclusions on the rationality of expectations: behavioural and psychology researchers overwhelmingly find various biases in subjective probability assessments, while most economists find evidence supporting rational expectations. The discrepancy arises from different understandings about rationality in assessing subjective probability. The psychology literature focuses on finding obvious violations of the fundamental principles of reasoning and judging under uncertainty, while the economic mainly focus on how people update their subjective assessments upon the arrival of new information. For example, Tversky and Kahneman [1974] described three heuristics that are employed in making judgments under uncertainty. Bernheim [1990] formally explored how the elderly form expectations by analyzing the response of self-reported expectation date of retirement to new information about Social Security benefits using the Retirement and Health Study data.

Instead of directly examining the process by which people form expectations, many literatures shift to evaluating subjective probability assessments or more broadly subjective judgments by various criteria. This approach focuses on what people really do, and generates insights about what people should do in an ideal environment. Roughly speaking, these literatures can be classified into four groups. The first mainly compares subjective judgments and outcomes in clinical studies. In his seminal book, Meehl [1954] examined 20 studies that compared clinical and statistical predictions and concluded that the judgment of a statistical model outperforms trained experts. Meehl's book motivated much other work in the field. The general conclusion is actuarial judgment generally does better than human judgment (see Grove *et al.* [2000] for a meta analysis). The second group focuses on comparing judgments on various fundamental statistics and the true statistics. The statistical concepts are usually mean, variance, independence and randomness. The main conclusion here is that because of cognitive limits, people are ill-equipped for assessing statistics. The third focuses on testing the rational expectations assumption in relation to updating expectations based on Bayes' law. Bernheim [1990] perhaps is the first paper examining the rational expectations assumption using micro data. Benítez-Silva and Dwyer [2005] tested the rationality of retirement expectations following Bernheim's framework.

The last group consists of some recent papers attempting to understand how people weight various information in forming subjective probability assessments. Hurd and McGarry [1995] studied how the elderly assess the probabilities of survival and found that the subjective probabilities "covary with other variables in the same way actual outcomes vary with the variables". Holden, McBride and Perozek [1997] examined how personal characteristics and health conditions influence subjective expectations of nursing home use and found that both men and women incorporate what is known about nursing home risks into their own subjective expectations. Both of these works, however, draw conclusions on the rationality of subjective probability assessments based solely on the signs of the coefficients that predict the subjective probability. For this reason, these works represent an essential departure from previous works by using real survey data and checking the sign of a covariate in both the regressions.

This paper expands the information weighting idea by exploiting a core implication of rational expectations. While these papers compared the signs, in this paper we compare the magnitudes. As Manski [2004]emphasized "the (rationality) assumption *per se* does not specify the expectations the persons hold. It asserts only that persons hold objective correct expectation conditional on the information they possess." Therefore, we impose a strict condition on rational expectations. The condition requires an identical weight of an covariate in projecting the subjective and objective probability under appropriate specifications and assumptions.

3.2.2 Definition

Assume $\mathbb{Y}$ is the interesting random event, *e.g.*, entering a nursing home in the future. Under the general cases where $\mathbb{Y}$ is a continuous random variable, we end up with a very complicated distribution function. To simplify the analysis, we consider only the simplest case of a dummy variable $\mathbb{Y}$, that is the realized values of $\mathbb{Y}$ can only be zero or one. Thus, the distribution function can be concisely expressed as the conditional expectation function and the conditional probability function: $E(\mathbb{Y}|\mathbb{X}) = Prob(\mathbb{Y}|\mathbb{X})$.

Consider the data generating process (DGP) described by an objective distribution function $\mathcal{G}(\mathbb{Y}|\mathbb{X},\mathbb{Z} : \alpha^*)$, where α^* is the parameter vector describing the relationship between random vectors $\mathbb{Y}$ and $\mathbb{X},\mathbb{Z}$. Without knowing the objective distribution function, a researcher observing only $\mathbb{X}$ specifies the objective conditional distribution function $\mathcal{F}(\mathbb{Y}|\mathbb{X} : \beta^*)$, where β^* is the parameter vectors characterizing the objective distribution of $\mathbb{Y}$ conditional on random vector $\mathbb{X}$. A rational individual, diagrammed below, observing the history and the random variables $\mathbb{X}$ and $\mathbb{Z}$ is interested in the data generating process and forms her subjective distribution function $\mathcal{G}(\mathbb{Y}|\mathbb{X},\mathbb{Z} : \alpha^o)$. Again, without knowing the subjective distribution function, the researcher specifies the subjective conditional distribution function $\mathcal{F}(\mathbb{Y}|\mathbb{X} : \beta^o)$, where β^o is the corresponding β^* in the subjective distribution function and characterizes the individual's belief on the distribution of $\mathbb{Y}$ conditional on observed random vectors $\mathbb{X}$.

$$\text{History} \xrightarrow[\text{X } \textit{and} \text{ Z}]{\textit{Observe Covariates}} \text{Expectation } \mathbb{Y}$$

In this scenario, rational expectations and expectation bias are defined as:

Definition. *Rational expectations is defined as* $\alpha^o = \alpha^*$.

With this definition, we propose

Lemma. *Under rational expectations,* $\beta^o = \beta^*$.

Proof. Under rational expectations, $\alpha^o = \alpha^*$. Thus, $\mathcal{G}(\mathbb{Y}|\mathbb{X},\mathbb{Z} : \alpha^*) = \mathcal{G}(\mathbb{Y},\mathbb{X}|\mathbb{Z} : \alpha^o)$. Define $\mathcal{M}(\mathbb{X}|\mathbb{Z} : \gamma^*)$ such that $\mathcal{G}(\mathbb{Y}|\mathbb{X},\mathbb{Z} : \alpha^*) = \mathcal{F}(\mathbb{Y}|\mathbb{X} :$

$\beta^*) * \mathcal{M}(\mathbb{X}|\mathbb{Z} : \gamma^*)$. Similarly for $\mathcal{M}(\mathbb{X}|\mathbb{Z} : \gamma^o)$. Under rational expectation, we also have $\gamma^o = \gamma^*$. Thus, under rational expectations, we have $\beta^o = \beta^*$. □

The lemma essentially states that parameters describing the relationship between $\mathbb{X}$ and $\mathbb{Y}$ should be identical regardless of the specification of $\mathcal{F}$ if individuals form rational expectations. Since the true structural function $\mathcal{G}$ is not known generally, the lemma offers a tool to evaluate rational expectations without imposing strong constraints on $\mathcal{F}$. Thus, we have

Definition. *Subjective expectation bias of $\mathbb{Y}$ in terms of $\mathbb{X}$ is defined as $\beta^o \neq \beta^*$.*

It should be first emphasized that the definition only focuses on the comparison between the *ex ante* objective and subjective distribution functions. While the *ex ante* subjective distribution can be elicited from a survey, the *ex ante* objective distribution is given by the DGP and not known outside of a controlled experimental environment. Nevertheless, the definition captures the idea that explanatory variables should have the same weight in projecting subjective probability as in the objective one.

The definition implies that conditional subjective distribution of $\mathbb{Y}$ on $\mathbb{X}$ and $\mathbb{Z}$, $Prob(\mathbb{Y} = 1|\mathbb{X}, \mathbb{Z})$, is identical with the true distribution function. Thus, the rationality of expectations requires optimal perception, which of course can be a result of evolution as many have argued. Note, however, the optimal perception is also the correct perception, where correctness indicates identical with the true distribution of the DGP. If perceptions were not correctly chosen, there would exist unexploited utility or profit-generating possibilities within a system. Furthermore, it also implies the conditional subjective probability function on the unobservable $\mathbb{Z}$, $Prob(\mathbb{Z}|\mathbb{X})$, is identical with the objective one. Since $Prob(\mathbb{Y} = 1|\mathbb{X}) = Prob(\mathbb{Y} = 1|\mathbb{X}, \mathbb{Z}) * Prob(\mathbb{Z}|\mathbb{X})$, rational expectations leads to the conclusion $\beta^o = \beta^*$.

The idea that a covariate used to form subjective expectations should be utilized the same way as it predicts the objective distribution of $\mathbb{Y}$ is a very natural expansion and development of the conventional rational expectations assumption, which requires the coincidence of subjective probability and objective probability conditional on information. Since objective probability is mostly unobservable, the canonical rational expectations assumption is principally limited as a solution concept in empirical work. By contrast, the proposed definition does not directly compare subjective and objective probability and thus can be potentially adopted in many research fields. As a compromise, however, the definition does require the presence of covariates $\mathbb{X}$.

The idea is not novel in social sciences. McClland and Bolger [1994] and Hurd and MaGarry [1995] analyzed how aged people form their longevity expectations by using the HRS data set and noticed that "most remarkable, however, is that they (subjective expectations) covary with other variables in the same way actual outcomes vary with the variables." The proposal of attribute substitution in judgment first offered by Tversky and Kahneman [1974] and then revised by Kahneman and Frederick [2002] shares almost the same opinion about how people should make judgment under uncertainty, though it highlights the abnormal phenomena of representativeness and availability biases. In a nut shell, attribute substitution occurs when people, in making judgment, assess a specified attribute by substituting it with another attribute, which comes more readily to mind, *e.g.*, heuristics and anchors. The identification of attribute substitution overwhelmingly depends on being capable of finding a substitution attribute, which can be elicited in experimental conditions but is difficult otherwise. In part, this makes many people suspect that the reported judgment bias is only a result of the way the statistical information was presented to respondents(Gigerenzer [1991] and Hoffrage *et al.* [2000]).

Our definition exploits the implication of the rational expectations assumption that people should use information correctly and accurately: they should form subjective probability the same way as the covariates predicts objective probability. The real DGP is not known to us, nor is its structural form. Nevertheless, the parameters characterizing the reduced form of objective and subjective distribution functions should be identical if the rational expectations assumption holds.

Bias is identified as the discrepancy between β^* and β^o. The identification of β^o is guaranteed if we know the subjective probability of $\mathbb{Y}$ and covariates $\mathbb{X}$, with the aid of the specification $E(\mathbb{Y}|\mathbb{X} : \beta^o)$. The subjective probability cannot be directly observed, but can be elicited. Since the 1990s, various surveys have asked interviewees about expectations, *e.g.*, the Health and Retirement Study, which this paper focuses on, and the 1997 cohort of National Longitudinal Survey of Youth. Similarly, the identification of β^* can be achieved once we know the conditional probability, which can be estimated from the realized values of $\mathbb{Y}$. Notice, however, even with the identification of these two parameters, identifying biases still need to assure that the parameters are comparable.

Last, though bias is stated as "in term of $\mathbb{X}$", the interpretation can be much more flexible in practice. More concretely, consider two predictors x_1 and x_2. Assume $\beta_1^o = \beta_1^*$ and $\beta_2^o \neq \beta_2^*$, therefore the source of the expectations bias is incorrectly internalizing information from x_2. Further assume x_1 is observable and x_2 is not observable in the data set and the two predictors are correlated $x_1 = \rho x_2 + \vartheta$. Under regular conditions, the probability limit of the estimate of β_1^o and β_1^* would be $\beta_1^o + \rho\beta_2^o$ and $\beta_1^* + \rho\beta_2^*$. Since $\beta_2^o \neq \beta_2^*$, the estimates of the coefficient x_1 in the objective and subjective probability function will not be equal. Without knowing x_2, we thus conclude there is expectation bias in terms of x_1. In this sense, identifying a bias in terms of any covariates can actually point to the same bias sources in some extreme cases [2]. According to this argument, we should not overly emphasize a specific predicator in explaining bias. Of course, one would suggest testing the equivalence of all the coefficients instead of only testing the coefficient of a specific covariate. We do not because of practical considerations. First, it is convenient to frame the discussion in terms of some specific covariates. Except in extreme cases, identifying and analyzing

[2] One example of an extreme case is the real expectation bias only exists for the unobservable covariate z and all other observable covariates xs are correlated with z. Then, depending on the specification of probability distribution equation, we would identify bias in terms of various x.

bias in term of a specific covariate could help to shed light in exploring the root causes of the bias. Second, testing of the equivalence of all coefficients demands stronger assumptions on the distribution of estimates.

3.2.3 Identification Assumption

The identification of expectation bias needs to compare β^o and β^*, which requires observing both the *ex ante* subjective and objective probabilities. Since the objective probability is not observable in most cases, β^* has to be inferred from the *ex post* realization of $\mathbb{Y}$, Y.

Replacing the *ex ante* objective probability with the *ex post* realization introduces interim random events that would affect the realization of $\mathbb{Y}$. The following diagram shows how realization proceeds, following the formation of expectations. As above, an individual forms expectations after observing outcomes and covariates $\mathbb{X}$ and $\mathbb{Z}$, where $\mathbb{X}$ is observable to researchers and the individual and $\mathbb{Z}$ is only known by the latter. Before the realization of $\mathbb{Y}$, Y, some observable random event W and unobservables ϵ can affect the realization of $\mathbb{Y}$ but not the *ex ante* objective probability $\mathbb{Y}$. Since we are not directly interested in W and ϵ, we group them as random variable ζ [3].

$$\text{History} \xrightarrow[\text{X } \textit{and} \text{ Z}]{\textit{Observe Covariates}} \text{Expectation } \mathbb{Y} \xrightarrow[\zeta = (W, \epsilon)]{\textit{Randomness}} Y(\text{Realization of } \mathbb{Y})$$

The above concern is central to the identification strategy. In the following analysis, we first propose a strong assumption and then a weak assumption for identification. With the introduction of interim random events, the identification becomes under what conditions the β^* can be consistently estimated by observing the realizations Y. Notice β^* characterizes the relationship between $\mathbb{Y}$ and $\mathbb{X}$ in $\mathcal{F}(\mathbb{Y}|\mathbb{X} : \beta^*)$. After introducing the interim events ζ, consistency requires $\beta^{'}$ be identical to β^*, where $\beta^{'}$ (γ) characterizes the relationship between Y and $\mathbb{X}$ (ζ) in $\mathcal{F}(Y|\mathbb{X}, \zeta : \beta^{'}, \gamma)$. This

[3] Of course, we could control W in the objective regression if it is observable. But it will not eliminate the consistency problem since ϵ is never observable.

identification requirement can be better understood in two steps. First, it requires the value of β^* be preserved from $\mathcal{F}(\mathbb{Y}|\mathbb{X} : \beta^*)$ to $\mathcal{F}(Y|\mathbb{X} : \beta^*)$, which is always true. Second, it further demands the equivalence of β from $\mathcal{F}(Y|\mathbb{X} : \beta^*)$ to $\mathcal{F}(Y|\mathbb{X}, \zeta : \beta^{'}, \gamma)$. Depending on the specification of $\mathcal{F}()$, different assumptions will be needed to preserve the value of β^*. To simply our analysis, we focus on linear models.

Assumption. *Orthogonality to interim information, i.e.,* $\mathbb{X}$ *is orthogonal to* ζ *or* $\mathbb{X}^T\zeta = 0$*.*

Orthogonality to interim information is a weaker statement than exogeneity. It allows endogeneity to the the extent that interim information is orthogonal to $\mathbb{X}$, the population characteristics. To understand it, let us consider the example where we are interested in the probability of an individual aged 40 entering in a nursing home facility at age 80. Between age 40 and 80, many health and social factors will affect the realization of the interested event. The orthogonality to interim information requires events happening between age 40 and 80 are unrelated with population characteristics at age 40. This is a strong assumption and is a sufficient condition for identification. We later relax it to a weak condition. With this assumption, we have in the linear case

Claim. *If the orthogonal to interim information assumption holds,* $\beta^{'} = \beta^*$*, where* $\beta^{'}$ *describes the relationship between* $\mathbb{X}$ *and the realization* Y *and* β^* *describes the relationship between* $\mathbb{X}$ *and objective probability* $\mathbb{Y}$*.*

The proof can be achieved as follows. Assume the model without any interim events is $Y = \mathbb{X}\beta^* + \chi$ and the model with the interim events is $Y = \mathbb{X}\beta^{'} + \zeta\gamma + \chi$. Obviously $\beta^* = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T Y {=} \beta^{'} + (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\chi + (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\chi$. So, $\beta^* = \beta^{'}$ as long as the orthogonal to interim information condition holds.

Now, we discuss the weak condition for identification. The orthogonal to interim information assumption is strong because it ignores the fact that individuals form expectations on interim events ζ while forming expectations on $\mathbb{Y}$. If so, the information included in ζ but correlated with $\mathbb{X}$, $E(\zeta|\mathbb{X})$, should have no impact on identification. Hence, the orthogonal assumption reduces to

Assumption. *Orthogonality to residual information, i.e.,* $\mathbb{X}$ *is orthogonal to* $\zeta - E(\zeta|\mathbb{X})$*, where* $E()$ *is ex ante subjective expectation.*

Similarly, we have

Lemma. *If the orthogonal to residual information assumption holds,* $\beta^o = \beta'$ *under rational expectation, where* β^o *describes the relationship between* $\mathbb{X}$ *and subjective probability* $\mathbb{Y}$ *and* β' *describes the relationship between* $\mathbb{X}$ *and the realization* Y*.*

The proof of the lemma parallels that of the above claim. Let the model of subjective expectation be $\mathbb{Y} = \mathbb{X}\beta^* + \eta$ and the objective expectation model with the interim events be $Y = \mathbb{X}\beta' + \zeta\gamma + \eta$. Under rational expectation, let $\zeta = \mathbb{X}\tau + \nu$, therefore $\mathbb{Y} = \mathbb{X}(\beta^* + \tau\gamma) + \eta$ and $Y = \mathbb{X}(\beta' + \tau\gamma) + \zeta - E(\zeta|\mathbb{X})\gamma + \eta$. Following the same argument in the above discussion, we have $\beta^* = \beta'$ if $\mathbb{X}$ is orthogonal to the residual information $\zeta - E(\zeta|\mathbb{X})$.

Unlike its strong pair, the orthogonal to residual information assumption only requires being uncorrelated with unexpected information. In the above example, the expected part of the interim events will have no impact on identification. What matters are the unexpected part of the events from age 40 to age 80, *e.g.*, health shocks arriving after age 40. If the shocks are orthogonal to population characteristics at age 40, it is still possible to successfully identify bias. Cases where the weak condition is violated do exist. Consider the negative impact of smoking, which was largely unknown by society in the 1960s. Assuming smoking behaviour is correlated with other individual characteristics, which seems acceptable, the weak assumption is violated if smoking happens at the interim period. The next paragraph offers more detailed discussion on threats to the assumption.

Threats to the Weak Assumption The assumption is crucial for identification. Should it be violated, it may negatively affect any conclusions on expectation bias. Thus, it is valuable to know what factors can potentially invalidate it.

Most of the new events $\mathbb{X}'$ of $\mathbb{X}$ would not be a threat to the weak orthogonality assumption. Usually, $\mathbb{X}'$ is generated conditional on the outcome of $\mathbb{X}$. The correlation will violate the strong assumption, but not necessary the weak one. Following the previous example, health status between age 40 and 80 can be correlated with health status at age 40, therefore the strong orthogonality assumption is violated. However, the weak orthogonality still holds if the unexpected shocks to health status between age 40 and 80 are not correlated with health status at age 40.

A potential threat to the orthogonality between $\mathbb{X}$ and ζ could be moral hazard. However, in the nursing home market, it is hard to justify since entering into a nursing home facility usually needs at least two certified ADLs from professionals. Further, a threat to the strong assumption does not threaten the weak assumption, which requires moral hazard behaviour to be unexpected. This seems even harder to justify. The major source of the correlation between $\mathbb{X}$ and $\zeta - E(\zeta|\mathbb{X})$ is unexpected public policies. Typically, to achieve some objectives, an interim policy is designed based on the statistics of covariate $\mathbb{X}$ across observations. If this policy is not expected by individuals and therefore not integrated into expectations, the unexpected policy design could introduce a correlation between the policy and $\mathbb{X}$, violating the weak assumption. In the background of long term care, if an unexpected interim public policy that is based on the attributes of seniors in our sample is implemented, it might be a concern from an identification perspective.

3.2.4 Model and Test Method

In linear models, the conditional probability function of $\mathcal{F}$ can be written as:

$$Prob(\mathbb{Y} = 1|\mathbb{X}) = G(\mathbb{X}\beta^o) \tag{3.1}$$

and

$$Prob(Y = 1|\mathbb{X}) = G(\mathbb{X}\beta^*) \tag{3.2}$$

Given the above functional forms, it is tempting to estimate βs simultaneously by using a bivariate probit model. However, the bivariate normal distribution of the error terms [4] is inappropriate as the range of subjective probability $Prob(\mathbb{Y} = 1|\mathbb{X})$ is the interval from 0 to 1, not the points 0 and 1 as in the bivariate Probit model case.

Estimating βs simultaneously is relatively efficient. This can actually be achieved by applying a seemingly unrelated estimation (SUR) method with the linear probability assumption of $G(\mathbb{X}\beta) = \mathbb{X}\beta$. As with all linear probability models, due to the assumed constant marginal effects of explaining variables, estimators can easily fall outside the unit interval.

The second choice is to estimate equation 3.1 and 3.2 separately. A probit model can be applied to estimate β^o. To estimate β^*, we adopt the fractional probit method proposed by Papke and Woodridge[1996]. The fractional probit method was originally proposed to deal with fractional dependent variables which range from 0 to 1 as well.

Comparing with other procedures like Berkon's minimum chi-square method advocated by Amemiya [1981], the fractional probit method is very attractive because it doesn't require any *ad hoc* transformations of the extreme values of 0 and 1. Recovering the regression function to get the probability can be easily handled as in a probit model. More precisely, the procedure maximize the following log-likelihood function

$$l(\beta) = Prob(\mathbb{Y} = 1) * log[G(\mathbb{X}\beta^*] + (1 - Prob(\mathbb{Y} = 1)) * log[1 - G(\mathbb{X}\beta^*)]$$

which is identical with the log-likelihood function of a probit model, except the range differs. Under some regular conditions, consistency and asymptotic normality can be easily derived.

To test the equivalence of the coefficients from the probit and fractional probit estimation, a similar idea to Hausman [1978] is applied. In the Hausman test, where one estimator is efficient and one is consistent, the covariance of the two coefficients asymptotically equals to the variance of the

[4] Actually, the bivariate probit model requires the potential error terms be bivariate normal, not the error terms we presented in equation 3.1 and 3.2.

efficient estimator. Without the efficiency hypothesis, we calculate the covariance of the coefficients, using the sandwich form proposed by White [1996][5]. We only report the results from the probit and fractional probit method. The linear SUR method generates similar results and thus is omitted.

3.3 Background, Data and Description Analysis

3.3.1 HRS and Expectation Question

The Health and Retirement Study (HRS) is a biennial, longitudinal survey of Americans, first conducted in 1992. It was originally designed to analyze retirement transition and thus initially only included the cohorts born between 1931 and 1941. In 1998, the HRS was merged with another survey, AHEAD, of individuals born in 1923 or before to cover all Americans over 50. In the following years, other sub samples were successively added to make it representative of this population. As one of the largest and most ambitious social science projects undertaken in the USA, the HRS is carefully designed to reflect both academic and policy interests in the area of aging and retirement planning. The comprehensiveness of the survey contents makes it suitable to answer a variety of different questions. Its core sections include demographics, physical health, household mobility, family structure, work history, disability, income, wealth, insurance, cognitive ability, expectation, and retirement plan*etc.*. Other experimental modules, including risk preference and parental wealth among others, are randomly distributed to a fraction of the respondents.

One unique feature of the survey makes it well suited for the purpose of studying subjective expectation bias. It asks the respondents about subjective expectations of future events, including longevity, living standard after retirement, working after retirement, leaving a bequest, and moving to a nursing home in five years *etc.*. Though questions about expectations have been asked in other earlier surveys, e.g., the Retirement History Study,

[5]For a detailed discussion on the test in Stata, see Weeise[1999].

the absence of carefully designed instruments makes the data from those questions limited in empirical work. Bernheim[1989, 1990] found the interpretation of expectations to be somewhat problematic: do people think of the mean, the mode, or the median when they are asked about their "expectation"? Instead of vaguely asking the respondents "expectation", the HRS is carefully designed to ask the respondents' subjective estimate of the probability of some categorical random variable being true, *e.g.*, the chance of living to 85 or the chance of working full-time after retirement. Another advantage of the expectation of a categorical random variable is that answering it may require less cognitive effort, which might be crucial since most surveys present no incentives for answering questions. Furthermore, the HRS strictly restricts the answers in numerical form ranging from 0 to 100 percent, which logically gives it a probability interpretation. Since disputes about subjective data easily arise, we thus describe the design of the expectation questions in detail.

To familiarize respondents with the basics of probability, the expectation section always begins with the following introduction:

> Next we would like to ask your opinion about how likely you think various events might be. When I ask a question I'd like for you to give me a number from 0 to 100, where '0' means that you think there is absolutely no chance, and '100' means that you think the event is absolutely sure to happen. For example, no one can ever be sure about tomorrow's weather, but if you think that rain is very unlikely tomorrow, you might say that there is a 10 percent chance of rain. If you think there is a very good chance that it will rain tomorrow, you might say that there is an 80 percent chance of rain.

Before continuing the survey, the respondents were given the opportunity to do a warm-up exercise:

> Let's try an example together and start with the weather. What do you think are the chances that it will rain or snow tomorrow?

To help the respondents better answer the questions, all questions are immediately followed with a graphic of an interval scaling from 00 to 100. It is marked with words "Absolutely No Chance" at 00 and "Absolutely Certain" at 100.

After completing the expectations questions, *e.g.*, chances of leaving any inheritance and of working for pay for sometime in the future, the respondents aged over 65 after the survey year 1998 were asked the question:

> What is the chance that you will move to a nursing home in the next five years?

which is immediately followed with a definition of a nursing home. The respondents' answer would reflect their assessment of their health status, financial status, insurance coverage, and other related information.

Besides the subjective probability, another dependent variable is the dummy variable of the realization of the interesting event. Because of the longitudinal nature of the survey, we can observe if the state of interest was realized, *i.e.*, they moved to a nursing home in five years.

Comparing the frequency of the realized events and the self-reported probability would be interesting, but we must first consider how to select the covariates x. There are two principles that we should be mindful of. First, the covariates should be obvious and known by the respondents. This is a natural requirement since rationality is conditional on the information the subject possesses. Second, the covariates should be predictive of the objective probabilities. The better the predictive power, the more accurate the estimates, which will be crucial later for testing the null equivalence hypothesis.

3.3.2 Choosing Covariates

The identification of expectation bias relies on comparing the values of covariates $\mathbb{X}$ in predicting expectations and realized events. Many factors can affect nursing home utilization. Demand side factors include the health status, household income, family structure, race and age of the individual. The supply side factors are the cost of nursing home care and local regulatory policies *etc.*. However, including supply side factors in our analysis is difficult because of the lack of sufficient variation. The RAND version of the HRS does not include data on state of residence. If it were available we might be able to integrate more supply side policy factors such as those that vary at the state level.

Health status can be measured in different ways. One is the subjective evaluation of the respondent. However, measurement error could contaminate the analysis. Thus, we focus on objective measures of health status. The first indicator is the health condition, which is the sum of the following factors: high blood, pressure, diabetes, cancer, lung disease, heart problem, stroke, psychological problem, and arthritis. The other health indicators are the activities of daily living (ADLs) and the instrumental activities of daily living (IADLs). ADL functions include bathing, walking, dressing,eating, and getting in/out of bed. IADL functions range from using a map, a calculator, and a telephone, managing money, taking medication, shopping for groceries, and cooking meals. Both ADLs and IADLs are good predictors of later long term care utilization and are widely used in policy underwriting and pricing. Some researches suspect that many elderly people are not fully informed of the prediction power of ADLs. Thus, testing the expectation bias in terms of ADLs and IADLs can shed some light on the debate.

Another health indicator that is highly related with nursing home utilization is the respondent's past nursing home stay experience. Past experience could be a good predictor about future use of nursing home facilities, but can individuals properly use this information?

Household income can influence nursing home care utilization because the costs are high and are not covered by general insurance policies, *e.g.*, Medicare or health insurance sponsored by employers. In practice, most seniors depend on Medicaid to pay for their nursing home bill. In order to establish Medicaid eligibility, the seniors must meet both the household income and non-housing asset requirements. For seniors with high income, Medicaid requires the seniors spend down their income below some level to qualify Medicaid. After that, Medicaid only pays the remaining bill if the spending down income is not enough to cover the cost. In this situation, if seniors anticipated a need for nursing home care, they would give assets above Medicaid limit to children, increasing current consumption. Thus, the measurement of assets may not be accurate because of manipulation. Thus, assets are excluded from the analysis. The results are not sensitive to the inclusion of various asset measurements, though. Years of schooling and race are also adopted as other measures of income and social status that affect nursing home utilization.

Family structure could affect nursing home care utilization either through its effects on reducing body functional disability or by providing a substitute to nursing home care. We use number of children as a measure of family structure. Another potential substitute to formal nursing home facility is cohabiting with a partner. In practice, many seniors, specially males, are cared for by their spouse instead of entering a nursing home.

Age is correlated with many unobservable health and mental indicators and is used for insurance policy underwriting. It is also the most accurately measured and naturally, respondents are fully aware of their age. It would be interesting to examine if the respondents can correctly internalize the age information into the expectation formation process.

Unlike other status variables, holding of long term care (LTC) insurance policies suffers from self selection. According to the standard theory of adverse selection in insurance market, those holding an insurance policy are high risk individuals. In a recent paper by Finkelstein and McGarry [2006], the authors demonstrate evidence of multidimensional private information in the long term care insurance market. With regard to subjective probability, we can check if those with the insurance policy have a higher subjective probability or a higher objective probability of entering a nursing home facility. It would give some useful hints for understanding the insurance market.

In sum, we have three health status measurements: health condition, ADLs and IADLs; a proxy health indicator of past nursing home stay experience, two family structure measurements, insurance policy holding indicator, income, age and race.

3.3.3 Descriptive Statistics

This section compares the objective and subjective probabilities over various factors discussed above after presenting the descriptive statistics on the variables of interest. When it does not give rise to confusion, the word "probability" is usually used as subjective probability and the word "chance" is used as objective probability.

Table 3.1 provides some descriptive statistics for the wave 4, wave 5 and wave 6 samples respectively. Wave 4 is the first year the subjective probability questions were asked in the way described above. The newest available wave is wave 9, which means the earliest wave that we have actual nursing home use data is wave 6. In the table, the objective probability is measured as the realization of the random event of interest, which takes the value 1 if a respondent ever entered a nursing home during the five years period and takes the value 0 otherwise. Since not all respondents are required to answer the subjective probability questions, we limit our sample to these who did answer it.

Columns $1-3$ report mean values and standard deviations for wave 6, Columns $4-6$ show the descriptive statistics for wave 5 and columns $7-9$ for wave 4. At each wave, we show the statistics for the whole sample, then males and females sequentially. The first two rows in the Table report the average objective probability and subjective probability. In all the columns, the average subjective probability is very close to the average objective probability, though the former is about 1% higher, which is not statistically significant. Females are usually better judges of probability. For males, the subjective probability is always statistically significantly higher than the objective probability. The overestimation is about 2%, and since the overall probability for a male senior is about 10%, the difference is about 20% in relative value. For females, the subjective probability is almost identical to the objective one. These statistics clearly show that the self-reported subjective probability actually mimics reality quite well. However, the closeness does not necessarily imply rationality: it only asserts that people hold objective correct expectations conditional on what information

Table 3.1: Descriptive Statistics

	(Wave 6)			(Wave 5)			(Wave 4)		
	All	Male	Female	All	Male	Female	All	Male	Female
Objective Probability	0.130	0.100	0.151	0.127	0.093	0.151	0.116	0.088	0.136
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.007)	(0.006)	(0.006)	(0.007)
Subjective Probability	0.139	0.129	0.147	0.142	0.127	0.152	0.124	0.116	0.130
	(0.003)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Age-65	9.369	8.936	9.675	9.082	8.698	9.351	8.825	8.397	9.130
	(0.108)	(0.145)	(0.123)	(0.102)	(0.137)	(0.106)	(0.094)	(0.120)	(0.107)
Income	0.433	0.540	0.358	0.410	0.501	0.346	0.375	0.461	0.313
	(0.011)	(0.017)	(0.009)	(0.009)	(0.015)	(0.007)	(0.008)	(0.013)	(0.007)
Years of Schooling	12.408	12.676	12.220	12.227	12.427	12.086	12.069	12.268	11.928
	(0.073)	(0.102)	(0.067)	(0.079)	(0.113)	(0.070)	(0.081)	(0.100)	(0.081)
Health Condition	2.087	2.050	2.114	1.949	1.904	1.981	1.844	1.835	1.851
	(0.018)	(0.029)	(0.023)	(0.019)	(0.025)	(0.025)	(0.021)	(0.027)	(0.026)
ADLs	0.315	0.230	0.376	0.332	0.239	0.397	0.356	0.279	0.411
	(0.011)	(0.013)	(0.015)	(0.014)	(0.014)	(0.018)	(0.010)	(0.016)	(0.015)
IADLs	0.306	0.215	0.371	0.285	0.188	0.353	0.313	0.234	0.370
	(0.011)	(0.013)	(0.013)	(0.012)	(0.012)	(0.015)	(0.011)	(0.013)	(0.014)
Nursing Home Expe.	0.027	0.022	0.030	0.019	0.016	0.022	0.018	0.014	0.021
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
LTC Insurance	0.139	0.139	0.139	0.123	0.124	0.122	0.113	0.118	0.110
	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.005)	(0.006)	(0.005)
Cohabitation	0.579	0.766	0.447	0.574	0.756	0.447	0.581	0.769	0.447
	(0.007)	(0.008)	(0.009)	(0.008)	(0.010)	(0.008)	(0.007)	(0.009)	(0.009)
Children	3.199	3.302	3.127	3.164	3.259	3.097	3.110	3.225	3.029
	(0.041)	(0.045)	(0.047)	(0.042)	(0.042)	(0.050)	(0.047)	(0.050)	(0.055)
Black	0.122	0.122	0.122	0.121	0.114	0.127	0.119	0.113	0.124
	(0.009)	(0.011)	(0.008)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
N	8115	3401	4711	7954	3334	4619	7980	3392	4584

Notes: The objective probability is the measured as the realization of the random event. The data source is Health and Retirement Study, Version 1, RAND Company.

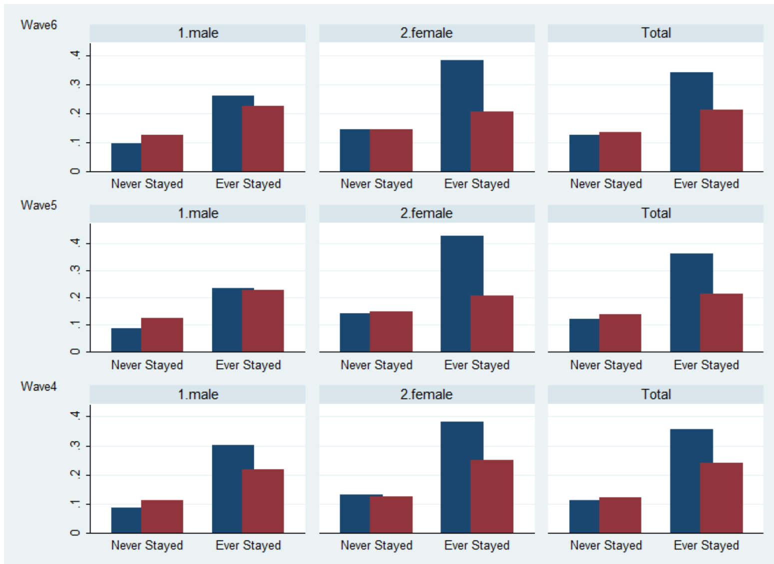
they possess. Without invoking the relationship between the probabilities and the information, it is impossible to evaluate the rationality assumption.

The above description might also give the impression that expectation bias exists only among males, not female seniors. This is not the whole story however. Figure 3.1 plots the average subjective and objective probability for different samples over the variable of past nursing home stay. The blue column corresponds to objecitve probability and the red column corresponds to subjecitve probability. The first panel corresponds to wave 6, the second to wave 5, and the third panel to wave 4. The figure shows that the chance of entering a nursing home is doubled if a male senior has previously used one, and the chance is tripled for a female senior. For males, the subjective probability follows the objective pattern, though with some noise. The striking feature is for the females, where they generally underestimate their chance of returning.

Figure 3.2 presents a similar plot over holding of long term care (LTC) insurance. Other research has indicated that those who choose to purchase insurance tend to be more at risk. This is not true in our figure, where these with a long term care insurance policy generally have the same chance to enter a nursing home as those without. What's more interesting in our figure is the pattern of the subjective probability: it is much higher than the true chance for male seniors with an insurance policy. It looks like the male seniors buy insurance policies because they think they are high risk people, not because they actually are. However, Finkelstein and McGarry [2005] showed that those with an insurance policy are more risk averse.

The similar chart over the age group is given in figure 3.3. Since age and gender are fundamental social characteristics that influence all kinds of decisions and judgments, it is natural to expect the subjective probability could better match the objective for a specific age group. Figure 3.3 does not support this idea. Actually, it indicates that the subjective probability is much higher than the objective chance for younger seniors. For female respondents aged over 80, the subjective probability is far below the true chance they will enter a nursing home. Interestingly, the female sample seems to support the idea that the self-reported subjective probability has

Figure 3.1: Comparing Objective and Subjective Probabilities: by Nursing Home Experience



some tendency to regress toward the mean: the subjective probability does increase with age, but less than the true chance.

Of course, the above figures only present suggestive hints about expectation bias and the rationality of expectations. This evidence is not conclusive since many other control variables are omitted in the figures. Specially, if the covariate and unobservables are correlated, any conclusion solely from the graphic analysis would be misleading without further evidence from formal regression analysis. Though the above statistics also indicate a satisfactory assessment of subjective probability, a formal evaluation on it can help shed light on some recent debates on the validity of the subjective measurements.

3.3.4 Assessing Subjective Probability

People are skeptical about subjective survey data, arguably with good reasons. Bertrand and Mullainathan [2001] summarized experimental evidence on how cognitive factors can affect how people answer survey questions. For

Figure 3.2: Comparing Objective and Subjective Probabilities: by Long Term Care Insurance Holding

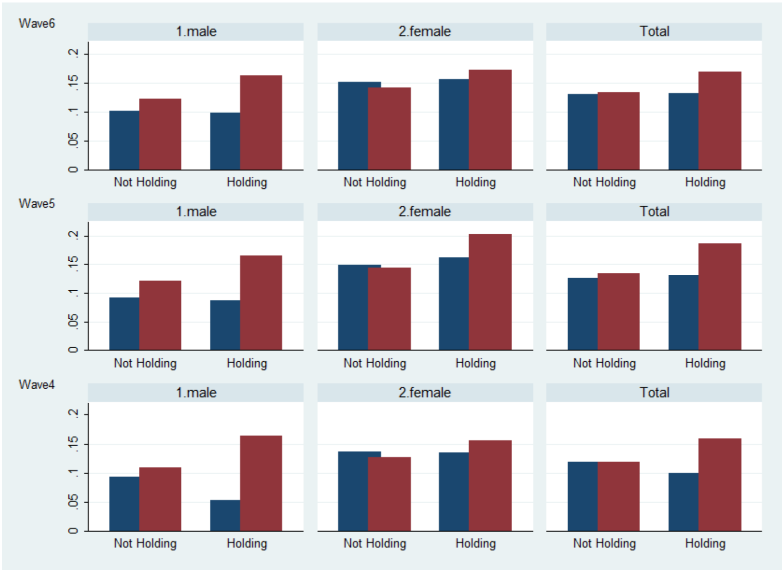
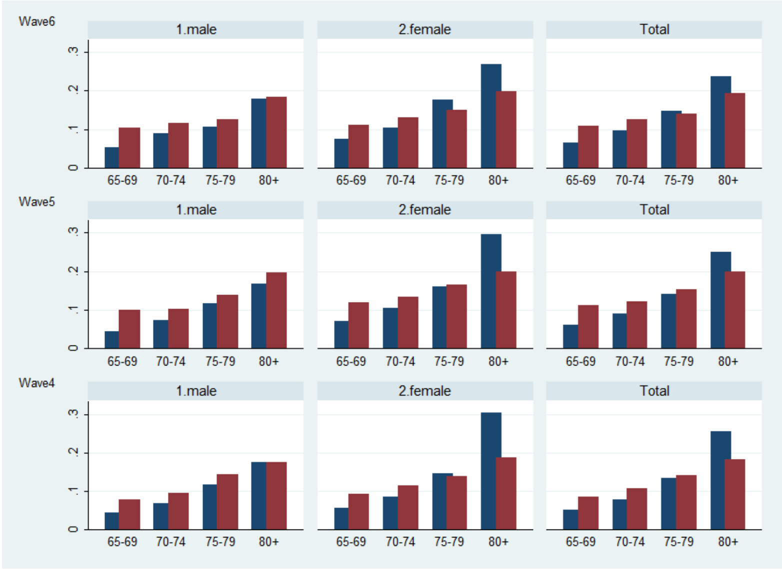


Figure 3.3: Comparing Objective and Subjective Probabilities: by Age



example, the answers can be manipulated by changing the order of the questions, modifying wording, or adjusting scale. Low correlations of the answers from two surveys spaced apart raise the doubt if the answers do measure something. Besides cognitive considerations, the objections against subjective data can be classified into two groups: doubts about if people are willing to truthfully reveal their subjective perceptions and doubts about if people have the ability to do so. Unwillingness to truthfully answer subjective questions can be partially a result of the absence of incentives and partially of the lack of verification. Though the absence of incentives is a common feature of all survey data, cognitive cost of answering makes subjective data particularly vulnerable. The extra cost arises because transforming beliefs and perception into numerical values is analytically challenging. As Tourangeau [1984] discussed, answering a survey question is a cognitive process involving comprehension of the question, retrieval and encoding of information from memory, assessment of the correspondence between the retrieved information and the requested information and finally communication. All these activities involve mental and cognitive cost. That there is no way to verify the subjective answers could lead to moral hazard in reporting the data, *i.e.*, reporting an easy but unrelated number in order to avoid the cognitive cost. The second doubts if people have the ability to do so. Mental cost, and the lack of cognitive ability evaluating the uncertainty in numerical form can impede validity. People would conjure up some figures if they could not express the perceived probability in numbers while are forced to do so.

In this section, we discuss some of these related arguments against self-reported subjective probability. Formally refuting all the assertions against subjective data is impossible. Instead, we focus on the objections mostly related to our data by examining the unwillingness and ability assertion. We achieve the goal by discussing some recent research supporting the application of subjective probability data. Following it is an assessment on how self-reported subjective probability can help to predict real nursing home utilization.

Evidence of good performance of using self-reported subjective probability in fitting data have been found by various researchers. Juster [1966]

Table 3.2: Can Subjective Probability Predict Behaviour?

	Wave 6			Wave 5			Wave 4		
	Subj. Prob.	Age	ADLs	Subj. Prob.	Age	ADLs	Subj. Prob.	Age	ADLs
Coefficient	0.13***	0.01***	0.07***	0.08***	0.01***	0.05***	0.05***	0.01***	0.03***
	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
R^2	0.01	0.04	0.04	0.01	0.01	0.01	0.01	0.01	0.01
Obs.	8295	8295	8295	6244	6244	6244	4892	4892	4892

Notes: The dependent variable is the dummy indicating the realization of the random event and the estimates are from probit estimation. Standard errors in parenthesis. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

found self-reported purchasing probability outperforms the verbal expression of buying intentions in explaining the automobile purchasing behaviour. Hurd and McGarry [2002] used the same data as our paper to examine how reliable the subjective survival probability is. They found the respondents actually modify their survival probability in response to new information and subjective survival probabilities do predict actual survival: those who survived in the survey panel reported about a 50% greater probability than those who died. As mentioned, Nyarko and Schotter [2002] and Bellemare, Krőger and Van Soest [2008] both reported better predictions using subjective probability data than merely assuming rational expectations in an experimental environment. In another paper exploring if seniors can understand the risk of moving to a nursing home, Taylor *et al.* [2005] found that the seniors reporting a higher probability of moving a nursing home within 5 years were indeed more likely to do so.

To assess how well the self-reported probability can predict true risks, we compare the predictive power of the subjective probability with ADLs and age, the covariates popularly applied in industry insurance underwriting. Table 3.2 lists the estimated coefficients in OLS regressions where the realization of entering nursing home is the dependent variable without controlling for other variables. Just as was found for the covariates age and ADLs status, the coefficients of subjective probability are always statistically significant across waves.

3.4 Age Bias and Income Bias

This section first presents the main results, then discusses some threats to the assumptions underlying the main results. In next section 3.5, we will also discuss the potential sources of expectation bias, which is different than the robustness analysis. The discussion on the robustness answers the concern whether the identified bias could be a result of a violation of the identification assumption or an inconsistent estimation of β^*, while the sources discussion seeks to find out what factors drive expectation bias.

3.4.1 Main Results

Tables 3.3 and 3.4 report regression results of equations 3.1 and 3.2 and the results of the Wald test for equivalence of the coefficients for various interesting covariates discussed above. All of the coefficients listed in the tables are the original estimates, rather than the marginal effects usually reported, as the magnitude of the coefficients is our focus. Wave 4 to Wave 6 results are separately reported, though mixing all waves generates similar conclusions. However, separating different waves data provides a cross validation and increases confidence in conclusions.

Table 3.3 presents the results from the male sample. Column 1 shows the estimates from a probit model where the dummy variable of realization is the dependent variable using wave 6 data, while column 4 uses wave 5 and column 7 uses wave 4. Standard errors are listed in parentheses. First, notice in most cases, we find the expected quantitative effect. The senior's age increases the probability of entering a nursing home, while education has no effect after controlling for income. Total income significantly decreases the chance of entering a nursing home. This might reflect that the rich choose private in-home care rather than a nursing home facility. ADLs have significant effect as well, but not IADLs and health conditions. This reflects the correlation between ADLs and IADLs and that ADLs measurements are better predictors of nursing home utilization. A health condition does not substantially affect the probability since it is a measure of autoimmune diseases that should be treated in a hospital. Past nursing home stay and

Table 3.3: Main Results: Male Sample

	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Age	0.033***	0.021***	5.145	0.030***	0.026***	2.693	0.0296***	0.030***	0.033
	(0.000)	(0.000)	(0.056)	(0.010)	(0.002)	(0.113)	(0.006)	(0.002)	(0.854)
Income	-0.253*	-0.014	5.835*	-0.188*	0.023	5.978*	-0.26*	0.023	5.185*
	(0.093)	(0.024)	(0.023)	(0.084)	(0.034)	(0.027)	(0.123)	(0.055)	(0.032)
Years of Schooling	-0.012	0.025***	4.01*	0.015	0.016*	0.336	0.021	0.016	0.095
	(0.014)	(0.003)	(0.051)	(0.017)	(0.010)	(0.577)	(0.025)	(0.016)	(0.766)
Health Condition	0.06	0.04*	0.71	0.09**	0.07***	0.44	0.06	0.05*	0.09
	(0.03)	(0.02)	(0.40)	(0.03)	(0.02)	(0.51)	(0.04)	(0.02)	(0.76)
ADLs	0.141**	0.073**	2.323	0.234***	0.095*	4.312*	0.194*	0.002	2.484
	(0.049)	(0.026)	(0.134)	(0.063)	(0.044)	(0.042)	(0.094)	(0.056)	(0.121)
IADLs	0.001	0.010	0.002	-0.015	0.046	0.195	-0.015	0.065	0.262
	(0.057)	(0.034)	(0.956)	(0.086)	(0.053)	(0.665)	(0.085)	(0.062)	(0.617)
Nursing Home Expe.	0.191	0.164	0.047	0.205	0.154	0.023	0.864*	0.265	2.138
	(0.172)	(0.113)	(0.857)	(0.254)	(0.159)	(0.880)	(0.374)	(0.223)	(0.155)
LTC Insurance	0.082	0.183***	1.099	0.10 6	0.232***	0.758	-0.123	0.255**	3.888*
	(0.082)	(0.043)	(0.306)	(0.125)	(0.063)	(0.396)	(0.176)	(0.082)	(0.054)
Cohabitation	-0.151*	-0.024	3.065	-0.068	-0.051	0.015	-0.158	0.033	2.453
	(0.072)	(0.044)	(0.092)	(0.097)	(0.065)	(0.923)	(0.117)	(0.065)	(0.132)
Children	-0.023	-0.042***	1.473	-0.013	-0.005	0.041	-0.003	-0.013	0.178
	(0.013)	(0.010)	(0.232)	(0.025)	(0.015)	(0.852)	(0.021)	(0.009)	(0.671)
Black	-0.092	0.134*	3.475	-0.213	0.084	2.586	-0.161	0.145	1.873
	(0.113)	(0.065)	(0.073)	(0.151)	(0.094)	(0.123)	(0.230)	(0.111)	(0.185)
Const.	-1.422***	-1.533***	0.313	-1.835***	-1.678***	0.444	-1.874***	-1.788***	0.086
	(0.161)	(0.084)	(0.585)	(0.193)	(0.101)	(0.513)	(0.197)	(0.104)	(0.781)
F Statistics	18.063	17.658	1.103	8.716	12.381	1.108	8.719	8.621	1.104
Obs.	3399	3399	3399	2562	2562	2562	2562	2562	2562

Notes: The Objective columns list the probit estimation results using the realization of the event as the dependent variable. The Subjective columns lit the fractional probit estimation results using the subjective probability as the dependent variables. The Test columns list the F-value of testing the equality of the proceeding two estimates. Standard errors are listed in the parenthesis in objective and subjective columns, and p values are listed in the parenthesis in test columns. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

Table 3.4: Main Results: Female Sample

	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Age	0.043***	0.022***	24.144***	0.053***	0.027***	25.895***	0.054***	0.027***	38.679 ***
	(0.006)	(0.003)	(0.002)	(0.004)	(0.005)	(0.004)	(0.001)	(0.004)	(0.001)
Income	-0.242*	0.033	6.124*	-0.002	0.070*	1.126	-0.091	0.044	1.754
	(0.103)	(0.034)	(0.021)	(0.061)	(0.037)	(0.295)	(0.087)	(0.055)	(0.195)
Years of Schooling	0.013	0.036***	2.276	0.025	0.017*	0.065	0.021	0.024***	0.586
	(0.011)	(0.012)	(0.144)	(0.016)	(0.010)	(0.805)	(0.013)	(0.013)	(0.457)
Health Condition	0.061**	0.062***	0.003	0.093***	0.063***	1.574	0.094***	0.076***	0.204
	(0.022)	(0.013)	(0.963)	(0.022)	(0.012)	(0.222)	(0.023)	(0.024)	(0.666)
ADLs	0.101***	0.022	7.505**	0.084**	0.044*	2.055	0.075*	0.054**	0.153
	(0.022)	(0.025)	(0.013)	(0.024)	(0.026)	(0.164)	(0.034)	(0.025)	(0.704)
IADLs	0.065*	0.095***	0.976	0.070*	0.050*	0.490	0.044	0.075***	0.527
	(0.038)	(0.028)	(0.337)	(0.037)	(0.028)	(0.498)	(0.034)	(0.025)	(0.476)
Nursing Home Expe.	0.474***	0.066	6.045*	0.495**	0.017	8.468**	0.456**	0.210*	1.444
	(0.135)	(0.116)	(0.026)	(0.153)	(0.108)	(0.015)	(0.145)	(0.088)	(0.245)
LTC Insurance	0.112	0.123**	0.021	0.116	0.223***	2.706	0.066	0.169**	0.848
	(0.073)	(0.044)	(0.895)	(0.060)	(0.040)	(0.116)	(0.096)	(0.067)	(0.369)
Cohabitation	-0.100	-0.014	2.428	-0.156**	-0.041	4.806*	-0.125*	-0.028	2.313
	(0.053)	(0.034)	(0.133)	(0.043)	(0.037)	(0.032)	(0.054)	(0.035)	(0.148)
Children	-0.013	-0.018	0.056	-0.014	-0.017	0.148	-0.026	-0.028*	0.006
	(0.012)	(0.014)	(0.831)	(0.015)	(0.011)	(0.712)	(0.017)	(0.013)	(0.977)
Black	-0.128	-0.028	1.301	-0.159	0.046	2.916	-0.323**	-0.155*	2.284
	(0.085)	(0.056)	(0.264)	(0.102)	(0.054)	(0.091)	(0.104)	(0.069)	(0.145)
Const.	-1.695***	-1.783***	0.476	-1.925***	-1.557***	3.154	-1.938***	-1.755***	0.767
	(0.183)	(0.084)	(0.024)	(0.194)	(0.087)	(0.085)	(0.186)	(0.096)	(0.396)
F Statistics	28.920	33.204	1.051	30.049	31.537	0.459	27.303	29.632	1.216
Obs.	4709	4709	4709	4617	4617	4617	4583	4583	4583

Notes: The Objective columns list the probit results using the realization of the event as the dependent variable. The Subjective columns lit the fractional probit results using the subjective probability as the dependent variables. The Test columns list the F-value of testing the equality of the proceeding two estimates. Standard errors are listed in the parenthesis in objective and subjective columns, and p values are listed in the parenthesis in test columns. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

holding a long term care insurance policy only have mild effects. If most of the male seniors are enrolled in the policy through employee group purchasing, the mild effect seems reasonable. Though it is popularly believed that male seniors living alone have greater demand for long term care, the coefficient of cohabitation shows it is not so true after controlling for other confounding variables. Having more children does not have a strong effect. Most of the estimates have the sign we expected in the descriptive analysis. Second, the estimates are very close across the three waves. The age coefficients are almost identical across the waves. This could be a good sign for the specification and the selection of the covariates.

Column 2 lists the estimates from the fractional probit estimator where subjective probability is the dependent variable for wave 6, while column 5 for wave 5 and column 8 for wave 4. Again, standard errors are listed in the parentheses. Some interesting features deserve emphasis. First, almost all estimates have the same sign as that of the corresponding ones in the objective probability regressions. Second, the estimates across the three columns are generally close. This is obvious for the estimates for age, health condition and LTC insurance policy holding. The stable pattern supports the conclusion that self-reported subjective probability can be seriously analyzed. Column 3 lists the F-value from testing the hypothesis that the coefficients of the corresponding covariate in column 1 and column 2 are equal. Similarly, column 6 for the coefficients in column 4 and 5, and column 9 for the coefficients in column 7and 8. Here listed in the parentheses are the p-values.

The first row shows the results of the age covariate. In both the objective and subjective regressions, the estimates are precisely estimated and are very close. And as a result, the equivalence hypothesis is rejected only in wave 6, but not in other waves. Since age is highly correlated with other covariates and unobservables, the results shows that the male seniors actually did well in internalizing age information into their subjective expectation.

The second row shows that the household income does not predict subjective probability as it does objective probability in terms of both magnitude and significance level. It suggests male respondents do not correctly internalize household income information. The subjective estimates have the opposite signs to the objective ones, and the hypothesis of zero cannot be rejected. However, income does predict the objective chances very well. The F-values of the equivalence hypothesis test are 5.83 for wave 6, 5.97 for wave 5, and 5.18 for wave 4, all of which lead to a rejection of the hypothesis at a significance level less than 5%. The significantly negative estimates of the subjective income coefficients could be a result of the negative selection resulting from Medicaid, which requires seniors to satisfy the income limit for qualifying for Medicaid eligibility. Alternatively, it could be that lower income seniors are generally less healthy and thus have a higher demand for nursing home care. It could also be that higher income seniors have more long term care options, like home health care and informal care from family members. Thus, the income bias can arise either because of the lack of Medicaid eligibility information or a misunderstanding of the effect of income. To further understand the Medicaid effect on nursing home utilization, we also controlled for non-housing assets and/or other assets measurement, but the results are not sensitive to these and therefore are omitted.

The third row gives the estimates of years of schooling. The estimates in the objective probability regressions indicate that schooling years is not a powerful predictor of the chance of entering a nursing home, while the estimates in the subjective probability regression show that seniors slightly misunderstand it, *e.g.*, the subjective coefficient in wave 6 is 0.025(0.003), which is much higher than the corresponding objective coefficient. Overall, the equivalence hypothesis is only rejected at wave 6 with a significance level of 10%. Since schooling years is generally believed to be highly correlated with other unobservable ability or cognitive factors, the results can be considered as a sign indicating no serious measurement error in the subjective probability. Both age and education information are correctly internalized and the hypotheses are usually not rejected, but the underlying explanations can be different. Age is a significant predictor, while education is not. It should be an easier task to abandon some less significant information while forming expectation. To correctly use a significant predictor demands much more rationality and cognitive cost: the equivalence of the age coefficients represents a precise process of collecting, assessing, judging and applying information. Thus, rows 1 and 2 suggest the male seniors are quite successful in tagging important information and applying it in forming expectations.

The fourth row indicates that the seniors utilize health conditions very well: the magnitude of the estimates in both objective and subjective probability regressions are close enough that a F-test of the equivalence hypothesis cannot be rejected at any reasonable significance level.

The fifth row gives the estimates for the ADLs covariate. The estimates and tests reveal some mild evidence for optimism bias: the size of the estimates, though generally significant, is much smaller than that of the estimates in the objective probability function. It seems male seniors actually take less account of the ADLs in projecting their subjective probability than its true weight in forming objective probabilities. In other words, the seniors knew the value of the information included in ADLs, but somehow underestimate its magnitude. However, the optimism is only mild. The hypothesis is only rejected in wave 5 because, in other waves, the objective coefficients are not precisely estimated.

The sixth row gives the estimate for IADLs. The estimates are very small relative to the standard errors and the equivalence hypothesis can not be rejected. The estimate of past nursing home stay experience is given in seventh row. As we have discussed, after controlling for other characteristics, the past nursing home experience actually does not significantly predict the chance.

The eighth row is especially interesting. The objective coefficients show that those with long term care policies do not have a significantly higher risk than those without. However, the subjective coefficients are all significant at 1% level at all waves. In other words, those male seniors with a policy actually significantly reported a higher subjective probability, even after controlling for other characteristics. As we mentioned, most of the policy holders are enrolled through workplace benefit. This result reveals some important potential bias though the hypotheses are only rejected in wave 4. Once again, the non-rejections merely reflect wide dispersion of the objective coefficients. If the sample excluded those purchasing policies in the retail market, the equivalence hypotheses are rejected. We leave it to further research to discuss how bias can effect health insurance market outcomes.

The next three rows show that cohabitation, quantity of children and race do not affect the chance of entering a nursing home nor the subjective probability of entering a nursing home. The last row gives the results for the constant term. In all the cases, the equivalence hypothesis can't be rejected.

Table 3.3 thus presents some strong evidence the expectation bias in terms of household income and some mild evidence supporting the biases of age, ADLs and schooling. The conclusion should be read cautiously, however. Any identified expectation bias can actually be a result of unobservable factors excluded from our specification, which is discussed in the last part of section 3.2.2.

The results for the female sample are listed in Table 3.4, which has the same structure as Table 3.3. Columns 1, 4 and 7 list the estimates of the objective probability regression and column 2, 5 and 8 correspond to the estimates of the subjective probability regression. Though Table 3.4 shares some features with Table 3.3, it differs in the following points. First, the estimated objective coefficients of the covariate age are much higher than the corresponding coefficients for the male sample, which implies that age actually has a stronger effect on females than on males. However, the subjective coefficients for females are almost identical with that of the male sample. As a result, we find very strong age bias for female seniors across all waves. All of the equivalence tests are rejected at a level less than 0.1%. Since age is one of the easiest factors to assess and is informative to all seniors, the female age bias represents serious optimism among female seniors. Of course, some possible threats to the identified age bias do exist: it could be the older seniors do not correctly express the subjective probability because of wishful thinking.

Second, contrary to the male result, household income does not significantly affect the chance of nursing home utilization. The subjective coefficients vary from year to year and are not significant. This could be a result of the lack of response to income information while calculating the subjective probability. Over all, no robust income bias was found among female seniors.

Third, past nursing home experience has better predictive power for females than for males. In all the estimations, the objective coefficients are about 0.45 and highly significant, but the subjective coefficients are not of similar magnitude and the tests of equivalence are mostly rejected. It seems there is significant optimism among female seniors. Notice the female seniors have the same pattern as the male seniors regarding the long term insurance status.

Combining the evidences from the two tables, we conclude a significant age bias among female seniors and income bias among male seniors [6]. Biases in terms of ADLs, LTC insurance status and past nursing home experience are only mild. Understanding the true source of bias and gender differences can help us make some progress in understanding how people form expectations. In the remainder of the paper, we consider some potential sources of the biases after examining several threats to the main results.

3.4.2 Threats to the Main Results

Two major threats could invalidate the discussed biases. One is the violation of the identification assumption that is the weak orthogonality condition. The long term nursing home care industry in the US was highly regulated and various public policies were proposed and implemented in the last decades. Different states adopted different approaches to regulate long term nursing home care, but commonalities between jurisdictions allow us to evaluate if the regulatory regime is truly a threat to the results. Another major threat is the inconsistent estimate of β^o resulting from the non-classical measurement error of subjective probability, which could be caused by cognitive factors or wishful thinking by the seniors.

[6] The gender bias difference is an interesting question. Though many potential reasons can be drawn by daily observation on gender difference, one of particular interest is that female's own income might be less important, but age could be more relevant in marriage and labour markets.

Public Regulation Policy

Unlike many other health care markets, the nursing home market was not completely competitive because of supply regulation policies. Many states featured a moratorium on new entry or expansion; beds would not be added without a certificate of need (CON). A nursing home provider could not enter the market or add new beds without demonstrating the need for more beds, either by the occupancy rate or the ratio of bed to aged population. It is reported that the percentage of applications denied of all CON applications was about 40% in 1990s (Harrington *et al.* [1997]). In such an extensively regulated market, it is natural to suspect that regulatory policies would include some subjectively unexpected residual information correlating with the demographic and social economical profiles of seniors. However, scrutinizing the regulation policies during our study periods shows this worry is likely unfounded.

First, the CON's legitimacy is granted by the National Health Planning and Resources Development Act of 1974 (Public Law 93-64). The goals of the act are to improve the health of residents, to increase the accessibility of quality of health service, to contain health care costs, and to prevent unnecessary duplication of health services. To achieve these goals, the act required state approval for all new construction or expansion of health facilities. States quickly realized CONs are a cost containment strategy to limit rapidly increasing Medicaid expenditure. The CON regulation should be within the framework of Public Law 93-64, which is openly available to the market and consumers alike. In other words, it is unlikely there is any subjectively unexpected part of the policies that would be systematically correlated with the market characteristics.

Second, though states implementing CON vary over years, it was stable during our study period except for Indiana, which repealed the CON law in 1999 again after reinstating it in 1997 following an initial sunset of the law in 1996. Note, survey wave 4 was conducted in 1998, wave 5 in 2000 and wave 6 in 2002. The fluid situation in Indiana might partially explain the observed rejection of the equivalence hypothesis of the constant term in wave 6 of male sample, but should has no impact on our major results. In all these years, though amendments are frequently made to the CON laws, there are no major amendments concerning the regulation of the nursing home facility.

Given these observations, it seems fair to claim that the policies regulating the LTC market during our study period do not include any subjectively unexpected information that is correlated with the controlled observables in the market.

Measurement Error

Another major threat is the possibility of inconsistent estimation of β^o on the subjective side due to the measurement error of subjective probability. Though there are potentially many factors affecting the measurement error, the cognitive factors are the major concern since estimating subjective probability is a process of applying cognitive ability to formulate a quantitative result.

This section tests whether there is non-classical measurement error due to cognitive ability in the variable of subjective probability. Our method exploits the assumption that if cognitive factors undesirably affect the measurement of the probability of entering a nursing home, it should affect measurement error of other subjective probabilities. Consider the reported subjective probability $p_s = p_s^* + \xi$, where p_s^* is the true subjective probability and ξ is the measurement error.The measurement error ξ includes all cognitive factors that influence how people report the subjective probability p_s. Though ξ is unobservable, we have encountered its proxy earlier in our discussion. In the warm-up exercise(See discussion in section 3.3.1), the respondents were asked to assess the chance of raining tomorrow. Similarly, the reported probability of rain can be decomposed as the true probability and the measurement error. Since the true probability of raining is random, it should be uncorrelated with other subjective probability judgments. Thus, regressing the subjective nursing home probability on the reported probability of rain would generate a zero coefficient in the absence of systematic

Table 3.5: Assessing the Measurement Error due to Cognitive Ability

Waves	4	5	6
The Probability of Raining	.00012(.134)	-.00005(.585)	-.00004(.595)
R^2	.0003	.0001	.0000
N	9359	9568	10118

Notes: The dependent variable is the subjective probability of entering a nursing home. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

effects of cognitive factors. Alternatively, if the zero hypothesis were rejected, cognitive factors may shape the measurement in an undesirable way.

Table 3.5 lists the regression estimates for the survey waves from 4 to 6. All of the estimated coefficients are very small and not significant. The extremely low values of R^2 support the claim as well. Thus, the evidence suggests that the measurement error due to cognitive factors cannot be a driving force of the identified biases.

Another interesting factor affecting measurement error is wishful thinking. Though the definition varies from case to case, it has been documented in many fields. Babad and Katz [1991] found soccer fans over estimated the chance of victory for the side they favoured. Gordon *et al.* [2005] showed in an experimental environment that people selectively remember and forget to bias toward the information sources that are consistent with desired outcomes.

Directly assessing the impact of wishing thinking on measurement error is difficult since the later is unobservable. Alternatively, we assess the impact of wishful thinking on female age bias by utilizing an implication of wishful thinking. The wishful thinking argument notices the age bias are mainly driven by the smaller coefficients of age in the subjective probability equations and suspects the subjective probability measurement error is positively correlated with age, that is the older respondents intentionally report a lower subjective probability because of wishful thinking [7]. If this

[7] Of course, lower should be read as lower than the true subjective probability.

Table 3.6: Is Age Bias a Result of Wishful Thinking?

	Panel A: Compare the Age Estimates By Age Groups					
	Wave 6		Wave 5		Wave 4	
	Older	Younger	Older	Younger	Older	Younger
Age	0.01**	0.02**	0.02**	0.02*	0.03***	0.02
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
F Statistics	28.92	32.15	23.97	23.99	23.54	31.45
Obs.	2270	2441	1732	1816	1247	1503

	Panel B: The Pattern of Age Coefficients		
	Wave 6	Wave 5	Wave 4
Age	0.01**	0.02**	0.02**
	(0.00)	(0.00)	(0.00)
F Statistics	33.25	36.52	31.31
Obs.	2761	2760	2747

Notes: Panel A lists the age estimates for the age group above the age median and the group below the age median of the subjective probability regression. Coefficients of other controls are omitted. Panel B lists the age coefficients of the subjective probability regression for the same respondents at various survey years. Coefficients of other controls are omitted. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

were true, the age bias would merely be a reported bias: the identified age bias is purely a result of the measurement error rather than not correctly internalizing information. To evaluate this statement, we divide the female sample into two groups: those above median age and those below median age groups. Should the theory hold, we will observe a smaller subjective coefficient in the above median age group.

Panel A of Table 3.6 presents the estimation results. Obviously, the estimated coefficient for the younger seniors is actually smaller than the older sample in wave 4, equal in wave 5, though is lower than that of older seniors in wave 6. The evidence implies that the age bias is not likely due to wishful thinking or intentionally underreporting of the subjective probability by the seniors.

As a complementary analysis, we also examine the age coefficients of the identical respondents at different ages. More concretely, the method compares the age coefficients of the subjective probability regression for the same respondents in various survey years. The various ages are available because of the panel structure of the data set. By restricting the sample available in all three waves, the data set ends up with 2761 respondents. The intuition here is also from the implication of wishful thinking: if the seniors did report a lower subjective probability while becoming older, the age coefficients should decrease from wave 4 to wave 6. In panel B of Table 3.6, though the estimated age coefficient does decrease from wave 5 to wave 6, the decrease is not observed from wave 4 to wave 5. The lack of a consistent pattern of the age estimations thus weighs against the wishful thinking conjecture.

3.5 Sources of Biases

The finding of age and income biases brings the interesting question about the root causes of the biases. A root cause or source is any factor that (i) influences the assessment of subjective probability but has no effect on objective probability and (ii) the hypothesis of equivalence of β^o and β^* of the biased covariate could not be rejected after controlling for the source in subjective regression. Consider the age bias of female seniors in the above analysis. In all three waves, these seniors assign a lower weight on age information in forming their subjective assessment. Various reasons can potentially cause the departure of subjective assessment from the *right*, objective weight. For example, if cognitive ability is the source of age bias, it should not affect the objective probability but increases age coefficient in the subjective regression so that the equivalence hypothesis of age coefficients would NOT be rejected if cognitive ability is controlled for in the subjective regression.

Any potential sources of bias should first be tested against condition (i), which especially requires exclusion of a potential source from objective probability. Two potential bias sources, cognitive ability and a misconception between risk aversion and probability, are believed to satisfy the condition.

While it is feasible to validate that both factors have substantial impact on subjective assessment, verifying or refuting the exclusion requirement is more or less subjective. By simply maintaining the exclusion assumption, the next subsections are devoted to demonstrating both potential sources' impact on subjective probability assessment and reporting the empirical results of hypothesis tests.

3.5.1 Cognitive Ability

It has been long known in the social sciences that subjective probability assessment depends on cognitive ability. Besides the seminal work of Tversky and Kahneman [1971, 1974, 1981, 1985], Hogarth [1975] formally discussed the implication of cognitive process for the assessments of subjective probability. He argued assessing probability is a "selective, sequential information processing system with limited capacity". Wallsten and Budescu [1983] wrote:

> Upon being asked to evaluate the probability of an outcome, a person will search his or her memory for relevant knowledge, combine it with the information at hand, and (presumably) provide the best judgment possible. That judgment will depend on what is retrieved from memory, what aspects of the current information are utilized, and possibly on the sequential order in which this is all integrated into a unified opinion.

Clearly, cognitive ability is believed to have a central role in the process of subjective probability assessment.

On the other hand side, age differences in various measures of cognitive function have often been reported(see Salthouse and Mitchell [1990] and Hertzog [1989]). These papers generally found that age differences are substantial and a considerable proportion of age related variance can be traced to cognitive ability limits.

These readings thus imply a potential bias source: the identified biases, especially age bias, might be attributed to cognitive ability limits in formulating the subjective probability. Fortunately, various measurements of

Table 3.7: Are Cognitive Factors the Source of Biases?

Panel A: Age Coefficient, Females Sample									
	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Age	0.04***	0.02***	19.95***	0.05***	0.02***	15.90***	0.06***	0.02***	15.79***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
F Statistics	28.92	25.62	1.97	23.99	31.02	1.97	13.44	24.97	1.97
Obs.	4709	4709	4709	3545	3545	3545	2747	2747	2747
Panel B: Income Coefficient, Male Sample									
	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Income	-0.25*	-0.01	5.77*	-0.18*	0.02	6.01*	-0.26*	0.00	4.57*
	(0.09)	(0.02)	(0.02)	(0.08)	(0.03)	(0.02)	(0.12)	(0.04)	(0.03)
F Statistics	18.06	19.35	1.97	8.71	7.61	1.97	4.33	6.95	1.97
Obs.	3399	3399	3399	2562	2562	2562	2012	2012	2012

Notes: This table lists the similar results as the main result table except the cognitive factors are also included in the subjective columns, the estimates of which and other controls are not listed. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

cognitive ability are included in the HRS data set and thus can be formally controlled for in our regression. Briefly, these measurements include memory, mental status, vocabulary for most of the respondents, and numerical and quantitative reasoning for a small sample. The memory functioning measurements include self-reported memory, immediate word recall tests, and delayed word recall test. The mental status is measured by various tests that assess knowledge, language and orientation. These measurements include the serial seven test, backwards counting, date naming, object naming, and naming the President and Vice President of the United States. Because of a very small sample for the quantitative test, we disregard it in the analysis.

Table 3.7 lists the results after controlling for the above cognitive factors in the subjective probability function. Doing so does not change the result. In panel A, the F-value of the hypothesis test in wave 6 decreases from about 24.14 to 19.95, but the changes do not alter the identification of age bias.

The inclusion of the cognitive ability factors also slightly decrease most the F-values, but not the significance level. Since cognitive factors could also be correlated with income, panel B presents the results after controlling for cognitive factors in the subjective regression equation for male sample. Similarly, controlling for the cognitive factors only has a minor effects on the estimation. The evidence presented here thus indicates that although cognitive factors do have some minor effects on age bias and income bias, they are not the driving sources of the observed biases.

3.5.2 Risk Aversion

Another potential source of biases is the difficulty in separating risk aversion from subjective probability, *i.e.*, the reported probability somehow measures risk aversion as well. If so, the reported subjective probability contains some information about personal risk preference. Many articles have reported that age and other individual characteristics are correlated with risk aversion. For example, by studying the relationship between age and risky asset holding of Canadian households, Morin and Suarez [1983] concluded that risk aversion increases with age. Sung and Hanna [1996] analyzed the response to risk tolerance questions and weakly conclude that risk tolerance decreases with age. However, Wang and Hanna [1997] reported that risk tolerance actually increase with age after controlling for other individual characteristics. On the other hand side, correlation between income and risk aversion has also been reported. Shaw [1996] argued that risk aversion can affect human capital investment. By developing a model of joint investment in financial wealth and human wealth, Shaw showed that human capital investment is an inverse function of risk aversion and reported empirical evidence for the positive correlation between wage growth and risk taking.

This section develops a measure to proxy for risk aversion using the information available in the HRS and test if risk aversion could be a source of bias. Various methods have been proposed to measure risk aversion in the HRS data set. First, the ratio of risky assets to total assets could be a measure of risk aversion. The risky asset is defined as the net value of stocks, mutual funds and investment trusts. Using the actual outcome data, we can directly measure the interesting parameter and thus minimize the measurement error. The disadvantage is that the directly observed outcomes are also an equilibrium result and thus include the idiosyncratic shocks and constraints. Alternatively, the income risk aversion test result measures risk aversion as well. In the risk aversion test, the respondents are asked to choose between a pair of jobs where one guarantees the current family income and the other one offers an even chance to increase income but also the chance to lose income [8]. The major disadvantage of responsive risk aversion is its subjective nature: its measurement is influenced by the same factors that influence subjective probability. The measurement also has only a small sample since it was only asked in wave 4. Considering all the benefits and costs, we thus focus on the behavioural measurement of risk aversion, that is the ratio of risky assets to total assets.

Table 3.8 shows the results of estimation after controlling for risk aversion in the subjective probability regression equation. Though doing so does affect the estimated coefficients, it doesn't change the general pattern of either age bias or income bias.

3.6 Conclusion

Recent academic consensus believes that incorporating subjective expectations into economic analysis of various decision making process can effectively improve prediction efficiency. To do so requires a better understanding of expectations formation.

This paper contributes to increasing academic interest in expectation formation by defining expectation bias and presenting a unified strategy of identification. The first innovation in our method is to exploit the implication of rational expectations, which says people form objectively correct expectations conditional on what information they possess. In the static

[8]More precisely, the risk aversion is measured at four levels by asking if the respondent would take a job with even chance of doubling income or cutting it by a half, a third or a fifth. The most risk aversion is defined as not accepting any of the above chance and choose the status quo.

Table 3.8: Is Risk Aversion the Source of Biases?

Panel A: Age Coefficient, Females Sample									
	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Age	0.04***	0.02***	22.90***	0.05***	0.02***	13.82***	0.06***	0.03***	15.43***
	0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
F Statistics	28.92	25.64	2.32	23.99	21.33	2.32	13.44	17.48	2.32
Obs.	4709	4299	4709	3545	3227	3227	2747	2519	2519
Panel B: Income Coefficient, Male Sample									
	Wave 6			Wave 5			Wave 4		
	Objective	Subjective	Test	Objective	Subjective	Test	Objective	Subjective	Test
Income	-0.25*	-0.01	5.80*	-0.18*	0.02	5.95*	-0.26*	0.02	5.08*
	(0.09)	(0.02)	(0.02)	(0.08)	(0.03)	(0.02)	(0.12)	(0.05)	(0.02)
F Statistics	18.06	16.77	2.32	8.71	6.58	2.32	4.33	4.91	2.32
Obs.	3399	3159	3159	2562	2403	2403	2012	1893	1893

Notes: This table lists the similar results as the main result table except the risk aversion is also included in the subjective columns, the estimates of which and other controls are not listed. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

context, rationality is equivalent to the statement that the information represented by some related covariates should be the same weight in predicting both the objective and subjective probability. Following this intuition, we define expectation bias as the inequality of the coefficients of a covariate in objective and subjective probability equations.

Since objective probability is usually unavailable, identifying bias therefore requires a proxy for it. In our analysis, we substitute the objective probability with the *ex post* realization of the interesting event, which introduces interim factors to influence the realization. However, our discussion shows that as long as the interim events satisfy a orthogonality condition, substitution does not contaminate the identification strategy.

Our empirical analysis using the Health and Retirement Study finds primary evidence supporting expectation biases among seniors. More specifically, we find female seniors fail at correctly integrating age information when they try to formulate their subjective probability of entering a nursing home within five years. A similar result is also found for male seniors where they fail to correctly internalize income information. Robustness analysis suggests the biases are not driven either by measurement error in subjective probabilities or a possible intermediate event, *e.g.*, public policies regulating nursing homes.

Though cognitive ability has a minor effect on identified bias, it is not the major source. Neither is risk aversion. Our result presents a challenge to the conventional practice of assuming objectively correct expectations. Though it has been recently discussed in various experimental environments, our paper relies on real data, rather than observations in a lab environment. The generality of the result means that it can be potentially applied to understand various consumer behaviours under uncertainty.

Chapter 4

Risk Aversion vs. Bequest Motive: How do Seniors Make Long Term Care Insurance Decisions?

4.1 Introduction

Understanding how seniors make long term care insurance (LTCI) purchasing decisions is of importance for both policy makers and researchers. In the USA, long term care (LTC) expense is the single largest financial risk facings seniors with an annual cost of more than $200 billion. Yet, only about 10% of seniors are covered by any LTCI policies. According to classical asymmetric information theory, the LTCI market is expected to observe a positive correlation between insurance coverage and incident occurrence due to adverse selection and moral hazard (See Fang, Keane and Silverman [2007] for a concise discussion). However, recent research has noticed that in the LTCI market those covered by an insurance policy do not experience a higher incidence rate than those who do not (See Finkelstein and MaGarry [2006] for a thorough discussion).

Many institutions and market factors might help to understand the issues of thin market size and the absence of positive correlation. Preferences, *e.g.* risk aversion and bequest motive, have been recently proposed as driving forces to explain these issues. For example, Pauly [1990] analyzed potential reasons for the small LTCI market size by arguing that a LTCI policy coverage's primary objective is to protect bequest, which are already excessive because of imperfect annuity markets. Regarding the absence of a positive correlation in the LTCI market, one explanation is that seniors vary in their risk aversion (see Finkelstein and MaGarry [2006]), in addition to exogenous risk status. In other words, it relies on a negative correlation between risk aversion and risk status to offset the potential positive correlation.

Naturally, risk aversion as well as risk status should play important roles in driving seniors' LTCI shopping decisions. Just as any other insurance purchasing decisions, a rational individual needs to evaluate her/his risk status (*e.g.* high *vs.* low) and decide to buy or not based on idiosyncratic risk preference. On the other hand, because of both the risk feature of long term care and practice of Medicaid policy, bequest motive would be an other potential driving force. For American seniors, the life time chance of utilizing long term care is substantial, which is estimated at more than 40%. Though the Medicaid program pays most of the LTC expenses, its estate recovery policy influences the decisions of seniors because the program actually seeks to recover all of the expenses it paid for a senior from estate upon the death of the senior.

Integrating a bequest motive into the model could facilitate better understanding of many phenomena related to seniors' consumption, saving and investment *etc.*. For example, Vidal-Meliá and Lejárraga-García [2006] tried to understand the annuities puzzle by introducing a bequest motive but failed to find it a significant factor influencing the demand for annuities at all. Specially, Kopczuk and Lupton [2007] examined the effect of observed and unobserved heterogeneity in the desire to die with positive net worth. They found that roughly three-fourths of the elderly single population has a bequest motive and concluded that among elderly single households, about four-fifths of their net wealth will be bequeathed and approximately half of this is due to a bequest motive. More recently, Lockwood [2011] disentangled precautionary and bequest motives by exploiting LTCI purchasing decisions and saving patterns across wealth distribution and found bequest motives very important in explaining dissaving behaviour and the small demand size of the LTCI market.

Based on these evidence, it seems promising to view a senior's LTCI decision by integrating both risk aversion and a bequest motive. To do so, this chapter develops a dynamic structural discrete choice model where a rational, risk averse and bequest motivated single senior has to decide at each period whether to buy a LTCI policy or not. In this model, both risk aversion and a bequest motive determine a senior's value function and therefore drive the senior's LTCI decision. By carefully constructing how seniors make insurance and consumption decisions, the chapter is capable of estimating a joint distribution of risk aversion and bequest motive from observed LTCI choices.

This chapter gives insights to many related debates over the LTCI market and Medicaid policy reform because of its estimate of primitive parameters, which are these directly describing preferences. First, it helps to resolve the positive correlation puzzle observed in LTCI market. Fang, Keane and Silerman [2008] discussed the advantageous selection in the Medigap insurance market, that is those with the Medigap coverage actually spend less than these without conditional on controls for the Medigap price, and concluded that cognitive ability is the sources of advantageous selection. Our paper complements their results in the LTCI market and finds that a bequest motive, rather than risk aversion, drives the absence. Second, though this chapter could not answer the thin market size issue directly, it does offer answers to a series of questions about how policy changes can affect the LTCI market size since our model captures relevant institutional details.

This rest of this article is organized as follows. In section 4.2, we present some introduction to Medicaid policy, especially its spending down and estate recovery policy. Since LTCI eligibility shapes our sample selection, it is also briefly discussed. Section 4.3 describes the data source and provides some primitive analysis on how seniors make LTCI decisions. Section 4.4 first provides a description on the structural model being followed by a discussion on identification. The identification depends on a monotonicity property of the structural model. Treatment on subjective beliefs and a parameter calibration method are also specified. The discussion ends with a description of the likelihood function and empirical estimation method. Section 4.5 presents our estimate of the joint distribution of risk aversion and bequest motive. We find substantial heterogeneity in bequest motive. The estimated risk aversion is quite homogeneous. By inferring each senior's idiosyncratic preferences *ex post*, we also find bequest motive is a major source for the absence of positive correlation. This section ends with analyses on various counterfactual policy effects on LTCI market size. Section 4.6 concludes.

4.2 Institution Background

4.2.1 The Eligibility Rules and Estate Recovery Policy of Medicaid

Long term care (LTC) [1] expense is the largest single health and financial risk facing seniors in the USA. The total expenses reaches 135 billion dollars in 2004 (CBO) and it is projected to double soon because of the retirement of the baby boom cohort. For individuals, the average annual nursing home care expense jumped from $34,156 in 1995 to $60,249 in 2004 (Stewart *et al.* [2009]). And a senior's lifetime chance of using long term care is more than 40% (Kemper *et al.* [1991]).

The funding of LTC expense comes from four sources. The majority is state Medicaid programs, which pay for roughly half of nursing home expense and about 70% of all bed-days. Another public insurance program, Medicare, pays about 20% of the expense, which is almost equivalent to the out-of-pocket payment. Private insurance accounts about 9% of the total expense (GAO, [2005]). Unlike Medicaid, Medicare funding is predominantly for post-acute care of short-stays following hospitalization. Thus, for seniors, the major sources are Medicaid and out-of-pocket expense.

However, Medicaid is a means-tested, needs-based social assistance program rather than a social insurance program. To build eligibility, individuals must meet certain functional criteria as well as state-specified income and

[1] Generally speaking, long term care includes nursing home care, home care and community based care *etc.*. In this chapter, we do not distinguish these types of care and always use long term care and nursing home care interchangably.

asset thresholds. The functional criteria generally require at least two "activities of daily living" (ADLs) limitations, which include bathing, dressing, eating, toileting, transferring (from a bed to a chair or vice versa), and walking across a room.

The details of income and asset tests vary from state to state, but there are some common features that should be taken into account when we model seniors' insurance decisions. On the income side, if an individual's income exceeds an income threshold published by the federal government, the individual is not eligible for Medicaid coverage. However, because of a spending down policy, or its counter part of Miller Trust [2] at some states, the individual could be eligible for Medicaid coverage after spending down his/her monthly income below the threshold. In either cases, the individual has to spend his/her income to the threshold level before Medicaid intervenes and pays the remaining balance [3]. On the asset side, the threshold varies a little from state to state, and in most states it is $2000 dollars for a single without dependents during our study period. In calculating qualified assets, all cash, saving accounts, stocks, IRA and other retirement accounts are accountable assets. The major excluded asset is real estate. Thus, the asset test requires an individual to spend the entire balance of non-house assets to the exemption level before becoming eligible for Medicaid assistance. In sum, the means test of Medicaid requires a senior to spend down most income and non-house assets to some thresholds before Medicaid pays the shortfall.

Besides the eligibility tests that could affect a senior's private insurance purchasing decision, the estate recovery policy of Medicaid is another rule that should be included in our model. Upon the death of a senior who ever benefited from Medicaid for nursing home care, a state must, by law, seek to recover all of the expense it made on behalf of the senior. It is called estate recovery because it is not executed until the death of the recipient. To avoid transferring assets by recipients upon death, states law puts pre-

[2] Miller Trust, or qualified income trust, can be used to qualify a Medicaid applicant with income in excess of the eligibility limit for long-term care assistance from Medicaid.

[3] Of course, if a senior's income is enough to cover all the monthly bill of nursing home care, Medicaid will pay nothing. So, theoretically, seniors with above average income have a stronger incentive to buy insurance.

death liens on recipients' home and other assets. Though an estate recovery is prohibited in certain instances when Federal law deems that the needs of certain relatives for estate assets take precedence over Medicaid claims, it is strictly executed if the recipient has no dependent relatives, which include surviving spouse and dependent children. Combining the eligibility and recovery policies, we can treat Medicaid as an interest-free need-capped credit lender without checking collateral.

4.2.2 Long Term Care Insurance Market

Despite the large risk, private LTCI market is relatively small. Though private insurance policy covers about seven million lives in 2007 (Senkewicz [2009]), the coverage is only about 10 percent of all population aged over 60. Unlike other acute health insurance markets where group purchasing bought most policies, almost 80% of LTCI policies are sold to individuals directly. The premiums charged for LTCI vary by the age when a policy was initially sold, with higher premiums charged to those bought at older ages. Thus, a senior will pay the same premium as what was paid at the first time as long as the policy was never lapse, subject to the adjustment of inflation and other cohort size effect [4]. Since the risk of using the benefit at younger ages is lower than that at older ages, the pricing of the LTC insurance is believed to reduce adverse selection at later stages. From the viewpoint of a potential policy buyer, the trade-off of buying today at a low premium rate or buying tomorrow at a higher rate always exists.

Beyond the premium trade-off, a delay of buying decision also risks a senior losing approval of her/his insurance policy application. As reported by McGarry and Finkelstein [2006], an LTCI application form generally includes detailed information on an applicant's age, sex, and membership in one of seven different health states defined by the number of limitations to instrumental activities of daily living (IADLs), the number of limitations to activities of daily living (ADLs), and the presence of cognitive impairments

[4] Adjustment to cohort size effect usually happens when risk incidence rate for a cohort is higher than expected by policy insurers and therefore the previous charged premiums are not enough to cover cost.

etc.. Any positive answer to these questions results in the application being denied. Though the disapproval risk will not be explicitly modeled in our analysis, it has strong implications for sample selection: all these seniors who could not get approved should be excluded from our analysis. Otherwise, inclusion of subjects who could not get covered by any private insurance policies will contaminate the inference of preferences.

In sum, the Medicaid eligibility rules and recovery policy are two important factors in modeling LTCI purchasing decisions, while the risk of being denied in LTCI policy applications alarms caution in sample selection.

4.3 Data and Primary Analysis

This chapter uses the Health and Retirement Survey (HRS) 2002 data set. The HRS is a biennial, longitudinal survey of American beginning in 1992. The subjects surveyed in 2002 were also interviewed in 2004 which enables us to observe the incidence of entering nursing home facilities. Like many other surveys, the HRS includes very detailed wealth, health and insurance coverage information. However, some unique features make it well suited for our study. It carefully designs and asks various subjective probabilities of entering a nursing home facility and living up to 75 years old *etc.*. Those subjective probabilities provide very rich private information about how forward looking seniors envision future events.

The sample included in this chapter is seniors meeting the following demographic, health and wealth criteria: are single, between 60 and 70 and thus are covered or expect to be covered by Medicare policy, do not have any ADL or IADL disabilities, and have real estate assets more than $60,000 and yearly income at least $20000. To simplify the analysis, we also exclude those with negative liquid assets. We end with 696 observations. Undoubtedly, the selected sample consists of quite healthy and wealthy single seniors. In order to exclude possible home care offered by cohabitants, the sample is restricted to single seniors. As we discussed in subsection 4.2.2, the chance of getting a LTCI approval is small for those with any ADL or IADL disabilities. With the same purpose of more precisely pinning down the preferences (risk aversion and bequest motive), the threshold of housing assets is set to be about one year of long term care cost. According to American Health Insurance Plan [2007], about 95% of new buyers in year 2005 have an income greater than $20000. The highly selected sample would potentially estimate a less heterogeneous preference distribution. Alternatively, we could relax the selective criteria to have a more representative sample, *e.g.*, removing the real estate assets or income requirements. By including single senior with less real estate assets, it trends to have a smaller mean value but a greater standard error of bequest motive than otherwise. However, the impact of removing income requirement on the mean values of the distribution is not clear, though which would lead a more heterogeneous preference distribution.

Table 4.1 lists mean and standard deviation of various variables by LTC insurance policy coverage status. Several features of the table are specially interesting. First, incidence rates of nursing home care or risk status measured in about five year horizon across the LTCI holding status are very close. Actually, the incidence rate is 0.046 for non-holders and 0.050 for policy holders and the hypothesis of identical incidence rates would not be rejected. This is not surprising as the absence of positive correlation between LTC insurance coverage and incidence rate has been documented in other research. Second, policy holders and non-holders are comparable with respect to many other observable characteristics. Though policy holders are marginally richer, and more likely to be female, they have almost the same Body Massive Index (BMI) level, self reported health level and out-of-pocket medical expenditures as non-holders. Last, policy holders are very close to non-holders in the risk aversion category, but have a higher level self-reported probability of leaving bequest and probability of entering a nursing home facility in five years. Intuitively, were risk aversion a major factor driving seniors' LTCI decisions, we should observe its difference across LTCI holders and non-holders. Instead, what we observed are the differences in bequest motive and subjective incidence probability. We believe an attempt to understand how seniors make LTC insurance purchasing decisions should trace these facts.

Table 4.1: Descriptive Statistics by LTC Insurance Purchase Status

	Non Buyer	Buyer
Incidence Rate of Nursing Home	0.046 (0.209)	0.050 (0.219)
Age	64.98 (3.19)	64.88 (3.13)
Gender	0.29 (0.46)	0.17 (0.38)
Income	54,777 (287,099)	78,962 (240,518)
BMI	27.08 (5.04)	27.41 (4.87)
Self Reported Health Level	2.48 (1.00)	2.37 (0.86)
Out of Pocket Medical Expense	3,874 (17,239)	3,512 (7,506)
House Value	163,555 (133,501)	207,200 (300,861)
Risk Aversion Category	4.86 (1.45)	4.38 (1.60)
Subjective Probability of Nursing Home Care	10.62 (18.89)	12.57 (21.11)
Subjective Bequest Probability	51.46 (41.86)	64.66 (37.64)
Observations	577	119

Notes: Self Reported Health Level is coded with the range from 1 for Excellent to 5 for Poor. Risk Aversion Category is coded with the range from 1 for the least risk averse to 5 for the most risk averse in the hypothetical income game described in section 4.5.2. Subjective probability of nursing home care is the self-reported probability of entering a nursing home in five years. Subjective bequest probability is the self-reported probability of leaving any bequests. Standard deviations are in parentheses.

Table 4.1 presents some hints of understanding how seniors make LTCI decisions. One way to resolve the puzzle is that individuals hold a second dimension of private information which operates in opposite direction to offset the positive correlation between risk status and policy coverage. In the table, though policy holders do not have a significant higher risk status (0.050 *vs.* 0.046), the holders do have some private information: the subjective probability of leaving a bequest is higher than non holders. It also suggests that risk aversion, the preference to risk, is unlikely to be the second dimension of information: the non holders actually have a slightly higher risk aversion level. In contrast, the subjective bequest probability shows a higher value among policy holders than among non-holders. To form a better perspective, table 4.2 lists the average bequest probability and risk aversion level by insurance coverage and risk incidence status. For example, the average bequest probability for these not covered by any LTC insurance and who have not incurred incidence is 51.79 with a standard deviation 41.80. As we have seen above, the average bequest probability is higher for the policy holders than the non-holders and the average risk aversion is almost identical. Specially, The LTC insurance holders without utilizing the policy actually have a bequest probability of 65.77 versus to 44.56 for the non holders who actually entered a LTC facility. The similar comparison of risk aversion is observed at 3.31 versus 3.67, which again is too close to offset a positive correlation.

Based on the primary evidence, we therefore propose that a bequest motive, just like the already discussed risk aversion, is another major force driving LTCI decisions and might even play a more substantial role than risk aversion. However, the primary statistics are built on proxy variables for bequest motive and risk aversion. The next section develops a model to infer the risk aversion and bequest motive distribution in order to better understand how seniors make LTCI decisions.

Table 4.2: Bequest Probability and Risk Aversion by Insurance Coverage and Risk Incidence Status

Insurance Coverage	*Buyers*	*Non-buyers*
	Panel A: Bequest Probability	
Risk Incidence		
YES	64.66(37.64)	44.56(43.22)
NO	65.77(37.04)	51.79(41.80)
	Panel B: Risk Aversion	
Risk Incidence		
YES	3.66(0.81)	3.67(0.71)
NO	3.31(1.07)	3.27(1.08)

Notes: Standard deviations are in parentheses.

4.4 Model: Specification, Identification and Estimation

We start by discussing a structural dynamic discrete choice model of a risk averse and bequest motivated senior. The model captures the relevant practice of Medicaid policy and the front-load pricing of the insurance policy discussed above. Though a senior in our model optimally decides both the discrete choice of buying a LTCI policy and the continuous consumption level, our identification utilizes only the information from discrete choices by exploiting a monotonicity property of the model. It does not fully extract information from consumption choices since the observation is usually poor. See French and Jones [2010] for a different framework where moment conditions are employed to facilitate identification.

Our identification hinges on the structure of the model, especially a monotonic relationship between LTCI decisions and preference parameters. During analysis, we first consider the case where individuals follow the monotonic decision rule: purchasing an insurance policy if and only if the value of doing so is greater than the value of not doing so, $V^b > V^n$. Similar monotonic decision rule method was applied in recent works of Cohen and Einav [2007] and Einav, Finkelstein and Schrimpf [2010] *etc.*. Besides its heavy computation, it does not absorb any shocks and thus potentially provides a poor fit to observed data set. Thus, we later introduce an unobservable type I extreme value distributed utility shocks following the work of Rust [1987]. In this scenario, the probability of purchasing an insurance policy is thus given by a logit probability model.

4.4.1 Specification

We model the behavior of a rational, forward looking, risk averse, bequest motivated, retired (or nearly retired) individual with an accumulated stock of liquid and housing asset, uncertainty about nursing home state and mortality risk, and time separable utility. A similar framework has been adapted to model labour supply (Eckstein and Wolpin [1989]) and annuity choices (Einav *et al.* [2010]) *etc.*. In this model, a senior is assumed to maximize the present value of utility over an unknown finite horizon by choosing consumption and whether or not to buy a long term care insurance policy. At each period, a senior can be in one the following five states depending on whether he/she is alive and his/her nursing home and private LTC insurance coverage status. Let M represents the demise status, N the status of nursing home and D the status of insurance coverage. Specially, $M = 1$ means the state of demise, $N = 1$ means the senior is under nursing home care, and $D = 1$ indicates that the senior is covered by some private LTCI policy. There are only five statuses because the demise status $M = 1$ is an absorbing state: once an individual is dead, the values of the other state variables don't matter. The other four states are thus a combination of the N and D conditional on $M = 0$, e.g., $N = 1$ and $D = 0$ indicates the senior is staying at a nursing home facility while not covered by any private insurance policy. Other state variables are the age of the senior at current period a, the liquid asset of the individual w, the housing asset of the individual h, the annual income y, and subjective risks λ. However, the subjective risks λ will not be implicitly appeared in our dynamic equations. Rather they are explicitly embedded in expectation functions. For more discussion, readers are referred to section 4.4.3.

At any age $a < A$, the senior will die next period with some probability and after reaching age A, the mortality risk is assured. Further, individuals obtain utility only from bequest they left if they die during the current period. Thus, the value function is given by

$$V(a, w, h, D, N, M = 1) = b_i(w + h) \tag{4.1}$$

the welfare from unconsumed wealth after death can be a result of strategic bequest motive (Bernheim, Shleifer and Summers [1985]) or altruism (Tomes [1981]). Though we are not interested in the structural interpretation of bequests *per se*, we will come back to discuss the relationship between *ex post* expected value of bequest motive and children. Note specially, the utility function is idiosyncratically specified by the subscript i. When it does not cause confusion, however, we omit all the subscripts i standing for the discussed agent.

Equation 4.1 also implies that the model distinguishes two different assets w and h. We do so because of the eligibility and recovery policies of Medicaid as discussed above. Introducing two different assets brings several issues. First, we need to specify how an asset can transform to another asset. Under the Medicaid policy, asset transforming can be attractive depending on the distribution of assets. To exclude complexity, we assume that individuals can not transform assets from one to another. More specifically, we allow the dynamics of liquid assets to depend on the consumption choice and make the housing asset fixed up to a real interest rate r. This assumption, plus an additivity assumption of utility, thus normalizes the utility from real estate assets to be zero. Second, the initial distribution of the two assets is exogenous in the model. This is equivalent to saying that the manipulation of the asset portfolio has already been done before retirement age. Admittedly, the exogeneity of initial distribution of two assets is a strong assumption since the distribution would depend on unobservable heterogeneity in risk aversion and bequest motive. A more satisfying model should explicitly specify a function between preferences and the distribution by either speculation or assumption. We did not choose to do so because of extra computation burden. However, we did examine how *ex post* risk aversion and bequest motive is related to the initial asset distribution and only found some mild correlation [5].

The uncertainties in the model come from the uncertainty about nursing home care [6] and mortality risk. More specifically, the duration of nursing home stays is assumed only to last one period and be identical for each individual [7]. However, we allow the expense of nursing home care B to be idiosyncratic but known by each individual *ex ante*. Restricting the duration of nursing home simplifies unnecessary analysis complication and alleviate computational burden. Otherwise, we need to further specify the transition matrix of consecutive nursing home events (which can be age and heath dependent). In addition, the restriction does not adversely affect the estimate of the interesting preference distribution.

Last, the model implicitly assumes all expectations in the model are those self reported by seniors, rather than those inferred using *ex post* realizations as the conventional rational expectations method. Two expectations, those of mortality and of entering a nursing home facility, are the major uncertainties in the model. By adopting self-reported subjective probabilities, we do not need to implicitly specify age or health status variables when it does not cause confusions since those variables should be integrated into subjective expectation by seniors. For more discussion, readers are once again referred to section 4.4.3.

The timing of the model is the following. At the beginning of each period, the mortality uncertainty is resolved and thus known. If a senior is

[5]More concretely, we define the distribution as the ratio of liquid asset to total asset. In an OLS regression of the ratio on *ex post* risk aversion and bequest motive, though the interesting coefficients are statistically significant, the R^2 value is merely 1.2%. However, keeping in mind that our estimates of risk aversion and bequest motive are based on the initial asset distribution, the statistic significance does not necessary invalidate our exogenous assumption.

[6]It implies that nursing home status is determined by exogenous factors like the difficulties of activities of daily living (ADLs) and other supply factors like availability of nursing home care beds. Though to what extent seniors can choose to enter in to a nursing home *ex post* is an empirical question, it is reasonable to claim that entering into a nursing home is a random event *ex ante* since the seniors have to figure out the risk while making the policy purchasing decision.

[7]The average duration in a nursing home is about 2.04 years (Society of Actuaries [2007]), which corresponds to one period in our data set.

alive, she/he has to decide the consumption level and whether to pay the insurance premium or not. Each paid premium covers the this period only. And the premium is only determined by the initially purchasing age.

An individual obtains utility from consumption while alive. However, the individual does not evaluate the hospital care while in a nursing home facility. In other words, the service offered by a nursing home facility does not enter into the utility function. It is not clear how this affects our preference estimate. On one hand, if individuals anticipated the welfare from a nursing home facility and took it into account in decision making, ignoring the utility might overstate the mean value of risk aversion but underestimate bequest motive parameter. Conversely, treating nursing home care as neutrals (not goods nor bads) ignores the ability of workers to offset long term care shocks by adjusting their consumption. This leads us to inconsistently estimate the consumption risk facing uninsured individuals, and thus the mean value of risk aversion and the bequest motive. On the other hand, maintaining this assumption allows us to shut down the moral hazard channel in our model: obtaining utility form nursing home care would present some incentive to a senior to choose to enter into a nursing home. Therefore, nursing home risk is only treated as a financial uncertainty. More specifically, for $D=0$ or $D=1$, the value function of an a years old senior with liquid asset w and housing asset h who is in a nursing home facility is given by

$$\begin{aligned}
&V(a,w,h,D,N=1,M=0)=\max_{c}[u(c)+ \qquad (4.2)\\
&\beta*[Prob(M'=0|N=1,M=0)*W(a',w',h',D'=0,N'=0,M'=0)+\\
&Prob(M'=1|N=1,M=0)*V(a',w',h',D',N',M'=1)]]
\end{aligned}$$

where β is the per-period discount rate, the superscript prime $'$ designates next period, $Prob(M'=0|N=1,M=0)$ and $Prob(M'=1|N=1,M=0)$ are the probability of death and living at the beginning of next period conditional on current nursing home status $N=1$. In the above two mortality probabilities, age is intendedly omitted since in our empirical analysis these probabilities are extracted from self-reported subjective mortality probabilities (see subsection 4.4.3).

And the value function after experiencing and leaving a nursing home in the past is given by

$$W(a, w, h, D, N = 0, M = 0) = \max_{c}\{u(c) + \beta EW(a', w', h', D' = 0, N' = 0, M')\} \tag{4.3}$$

which is subject to a transformation condition of w' and h'. The constraint on $N' = 0$ reflects the assumption that a senior can not enter in to a nursing home again after being dispatched. Since $N' = 0$ holds for the future periods, there is no need to pay the insurance premium to get covered, $D' = 0$. This equation explicitly assumes all seniors at most have one chance to enter a nursing home facility. In other words, once they left a nursing home (if they did so before death), they either die or stay home next period. In practice, a typical senior usually visits nursing homes several times in her/his life time. However, multiple visits are irrelevant in our model since for a senior to make the purchasing decision what matters is the expected expense of long term care. Notice also we use value function $W()$ to denote dynamics after experiencing a nursing home here.

Equation 4.2 reflects the assumption that the nursing home care brings no welfare and the only utility sources are either from current and future consumption, or the bequest. In this equation, it implicitly allows a senior to choose an optimal consumption level to maximize utility level even when the senior is in a nursing home. In practice, seniors in a nursing home often sadly found they were aboard a boat preventing such optimization. However, completely removing such optimization in our analysis demands an alternative constrained optimization, which is even harder to justify.

We have not yet determined the dynamics of $w'$and h', which depend on both the insurance coverage and Medicaid policy. Consider the case where $D = 1$. Since a private insurance policy covers all the expenses of care, we have

$$w' = (w + y - c) * (1 + r) \tag{4.4}$$
$$h' = h * (1 + r) \tag{4.5}$$

where r is the per-period real interest rate. In this value function, we ignore the co-payment and deductable the individual needs to pay.

Consider the $D = 0$ case now. First, if the senior dies at the beginning of next period $M' = 1$, the Medicaid recovery and spending down policies indicate

$$w' + h' = max\{0, w + h + y - c - B\} * (1 + r) \tag{4.6}$$

where B is the known, idiosyncratic nursing home care expense. Second, if the senior survived after the nursing home care, that is $M' = 0$, then the dynamics are given by

$$w' = max\{2000, w + y - c - B\} * (1 + r) \tag{4.7}$$
$$h' = h * (1 + r) \tag{4.8}$$

The w' transformation reflects the Medicaid spending down policy in terms of the liquid asset and 2000 is the Medicaid allowance for each single senior that can be exempted from the spending down policy. The dynamics of w'' and h'' depends on the M''. If the senior is still alive at age a'', the dynamics of w'' and h'' are identical as the dynamics of $w'$and h' in equation 4.7 and 4.8. However, if $M'' = 1$, the estate recovery policy is executed and the dynamics become

$$h'' + w'' = max\{0, (w' + h') * (1 + r) + min\{0, w + y - B\}\} \tag{4.9}$$

where the inner min calculates the balance Medicaid paid on behalf of the recipient after the recipient spends down all the liquid asset. And $w+y-B > 0$, that is the individual is rich enough to pay the nursing home care bill out-of-pocket, then Medicaid paid nothing and can not execute the recovery policy. The outer max is obvious. Strictly speaking, the states indicating whether a senior ever in a nursing home should also be a state variable since it affects the later state space of N after a senior leaves the care facility.

If a senior is not in a nursing home, which is $N = 0$, at the beginning of the period, the senior decides whether or not to buy a private LTCI policy in additional to the usual optimal consumption. If the senior is not covered by any insurance policy $D = 0$, he/she can choose to buy a new policy and pay the premium determined by current age a. If not covered by any insurance policy $D = 1$, he can choose either to let the policy lapse and not pay any premium or continue paying the same premium as last period $p(a - 2)$. In this period, the senior faces a trade-off between buying an insurance policy this period to prevent the financial risk of next period and not buying a policy to maintain a higher expendable liquid asset. Moreover, for $D = 0$ or $D = 1$, the senior's value function is given by

$$\begin{aligned} &V(a, w, h, D, N = 0, M = 0) = \\ &max\{V^b(a, w, h, D, N = 0, M = 0), V^n(a, w, h, D, N = 0, M = 0)\} \end{aligned} \tag{4.10}$$

where $V^b()$ and $V^n()$ stand for the value of current states if the senior decides to buy a policy and not to buy a policy respectively, and are given by

$$\begin{aligned} &V^b(a, w, h, D, N = 0, M = 0) = \\ &\qquad \max_c\{u_i(c) + \beta EV(a', w', h', D' = 1, N', M')\} \end{aligned} \tag{4.11}$$

subject to

$$\begin{aligned} w' &= (w + y - c - p(a)) \geq 0 \quad if \quad D = 0 \\ w' &= (w + y - c - p(a - 2) \geq 0 \quad if \quad D = 1 \end{aligned}$$

and

$$\begin{aligned} &V^n(a, w, h, D, N = 0, M = 0) = \\ &\qquad \max_c\{u_i(c) + \beta EV(a', w', h', D' = 0, N', M')\} \end{aligned} \tag{4.12}$$

subjective to

$$w' = (w + y - c)(1 + r) \geq 0$$

where $a - 2$ is the age at last period [8], and $p(a)$ is the premium the senior pays. We maintain the standard assumption that seniors can not borrow against future income with the presence of mortality uncertainty. The expectation E is integrated over the distribution of the states of next period nursing home, N'. If the senior decides to buy a policy, then next period the senior is covered by an insurance policy $D' = 1$. Otherwise, $D' = 0$. The terminal condition is given by $V(A, w, h, D, N, M) = b_i(w + h)$ since demise is a sure thing after age A.

Since government transfers are generally available to bridge the gap between an individual's expendable income and the minimum consumption floor, the individual's consumption can not be less than some $\underline{c}$. Specially, we assume that the consumption level c automatically jumps to $\underline{c}$ once the optimal level is below the floor.

Finally, to recover the joint risk aversion and bequest motive distribution, we allow heterogeneity in both preferences. Concretely, the consumption utility function is assumed to be a constant relative risk averse (CRRA) function

$$u_i(c) = \frac{c^{1-\gamma_i} - 1}{1 - \gamma_i} \tag{4.13}$$

and the bequest utility function is assumed to be

$$b_i(w) = \delta_i \frac{(\underline{c} + w)^{1-\gamma_i} - 1}{1 - \gamma_i} \tag{4.14}$$

where c and w are consumption and bequest, γ_i and δ_i are the risk aversion and bequest motives parameters. Recovering the distribution of the two

[8]Notice that one period corresponds to two years in the model. In practice, the premium paid, conditional on the coverage history, depends on the age when the elderly initially bought the policy. The insurance companies frequently adjust the cohort premiums, though they can not charge a higher premium only because of aging. Note that $p(a-2)$ can be either the premium paid by a senior (which is determined by the age when the policy was initially bought) or the premium charged it it is sold at age $a - 2$.

parameters is the main task of the paper. Notice $\underline{c}$ is the minimum consumption level below which no bequest will be left and thus determines the extent to which bequests are luxury goods. In a static model with the above preferences, the optimal bequest level is a fraction of the after minimum consumption $w - \underline{c}$. Similar bequest preferences can be found in De Nardi *et al.* [2010] and Lockwood [2011] *etc.*.

In this model, both risk aversion and the bequest motive have the nice feature of monotonicity [9]: a larger value of γ_i implies a stronger risk aversion and thus a stronger desire to purchase an insurance policy while a greater value of δ_i implies a stronger bequest motive and thus a stronger desire to buy an insurance policy. Intuitively, all else equal, we expect that a senior with higher bequest motive would be more likely to buy an insurance policy. Since the risk aversion parameters appear in both consumption and bequest utility function, a higher level of risk tolerance makes insurance more desirable. In other words, observing the insurance decision helps us to infer both the value of the risk aversion parameter γ_i and the bequest motive δ_i.

Several features of the model need to be mentioned. First, we only model the uncertainties on demise and nursing home status of next period. All uncertainties on income y and the policy premium $p(a)$ are omitted from the model. It is a natural choice since we are mainly interested in the discrete choice of buying insurance, not the volatility of consumption. Besides, income y of seniors consists of social security and pensions *etc.*, which do not depend on employment status. Second, according to our model, switch of insurance coverage is mainly driven by the change in the perceived risks. In reality, as showed recently by Knotzka and Luo [2010], most of the policy lapses are driven by deterioration of financial status. To capture financial uncertainty, we could alter the model with a random income shock y, instead of fixing it. We choose not to do so because of the attempt to keep the model simple. Third, this is a partial equilibrium model where the

[9] The monotonicity feature actually is quiet broad. In includes many preferences that follow the specified utility functions. See Lockwood [2011] for a discussion on the bequest utility function.

pricing strategy $p(a)$ and the Medicaid policies are determined outside the model. A fully equilibrium model would be very attractive for answering some general equilibrium effects like how Medicaid eligibility rule changes can affect the welfare of seniors from different social groups. Last, other decisions that are observable in the data set and help to infer preference information, *e.g.*, life annuity purchases and choices in the Medigap market, are not integrated in our model. Doing so provides more vehicles to infer preference, but the estimated preference parameters will be entangled by various decisions.

The observation of seniors' choices plus the rationality assumption that the choices are made optimally according to the specified model provides information to retrieve underlying preferences. Intuitively, everything else equal, buying an insurance policy is optimal if and only if a senior is risk averse enough; and it is also true if and only if he/she has a strong enough bequest motive.

4.4.2 Preference Identification

There are two types of parameters in our model: one type about preferences γ and δ and another type about beliefs: nursing home expenditure B and subjective probability assessment of mortality and long term care risk. To estimate the model, we separately estimate these parameters following Rust and Phelan [1997] and Gourinchas and Parker [2002] *etc.*. Unlike others, beliefs in our data are partially observable thus our focus is on the preference parameters. This subsection discusses the identification of preferences that is followed by the discussion on beliefs.

To understand how to achieve identification, we first consider the case where seniors making LTCI purchase decisions strictly follow a monotonic rule, that is $d = 1$ if and only if $V^b > V^n$. Without any utility value shocks, buying is optimal if and only if the risk aversion preference is high enough, and for these currently covered by some private policies, not buying is optimal if and only if the risk aversion preference is low enough, everything else equal. Similarly, for these not covered by any private policies, buying is optimal if and only if the bequest motive is strong enough, and for these currently covered by some private policies, not buying is optimal if and only if the bequest motive is low enough, everything else equal.

To better understand the intuition, we draw a figure which shows the area where it is optimal for a senior to buy an insurance policy. In figure 4.1, the horizontal line represents bequest motive while the vertical line represents risk aversion. First, the monotonicity of the decision regarding risk aversion can be intuitively understood that there is a cutoff value of $\gamma_i^*(\delta_i)$ such that, holding the bequest motive constant, the senior will choose $d_i = 1$ if and only if the value of $\gamma_i \geq \gamma_i^*(\delta_i)$. Second, a similar conclusion can be made for the bequest motive parameter. Combining the above results gives an indifference curve, which is the locus of all the γ_i and δ_i such that, conditional on the current status D, buying an insurance policy brings the senior the same state value as choosing not to do so. Further, under any value combination of (γ_i, δ_i) above the indifference curve, the senior will choose to buy a policy $d_i = 1$ and under any combination below the indifference curve, the senior will choose not to buy $d_i = 0$. Thus, the area below the indifference curve and the two axes gives the probability that the senior will choose not to buy a policy at the cumulative distribution function $F(\gamma, \delta)$. Since we already expressed the probability of not buying a policy as a function of the joint distribution of risk aversion and bequest motive, we can recover the joint distribution function by maximizing likelihood based on the decision choices.

To formally prove the identification idea, we introduce some notation first. The observed data on each individual i who is alive $M_i = 0$ and not in a nursing home facility $N_i = 0$ are $(D_i, d_i, w_i, h_i, \lambda_i, x_i)$, where D_i is observed insurance coverage states, d_i is the insurance policy purchasing decision which is 1 if the individual decides to buy a policy and 0 otherwise, w_i and h_i are the liquid and housing asset respectively, λ_i is the subjective risk reported by the senior and x_i is a vector of observed characteristics. Identification is achieved if and only if the joint distribution function $F(\gamma, \delta)$ is uniquely recovered given enough observations.

The identification requires the following three assumptions:

Figure 4.1: Optimal Decision Without Utility Shocks

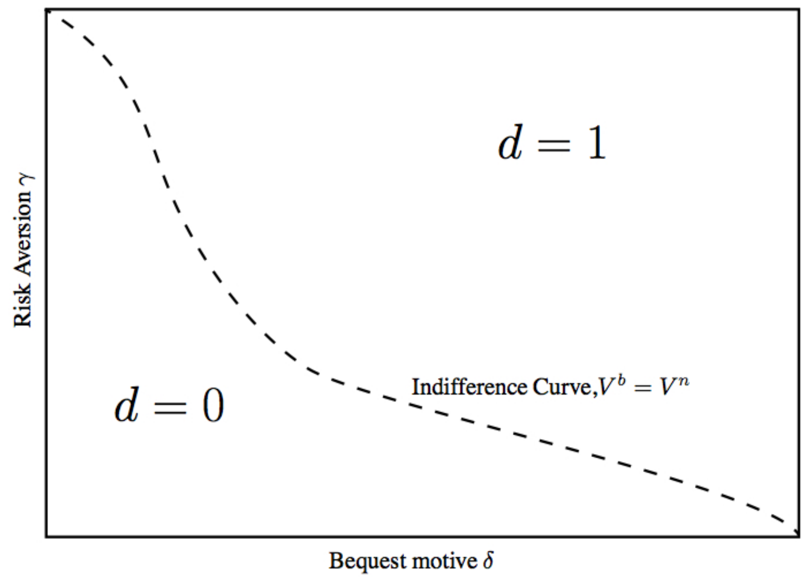
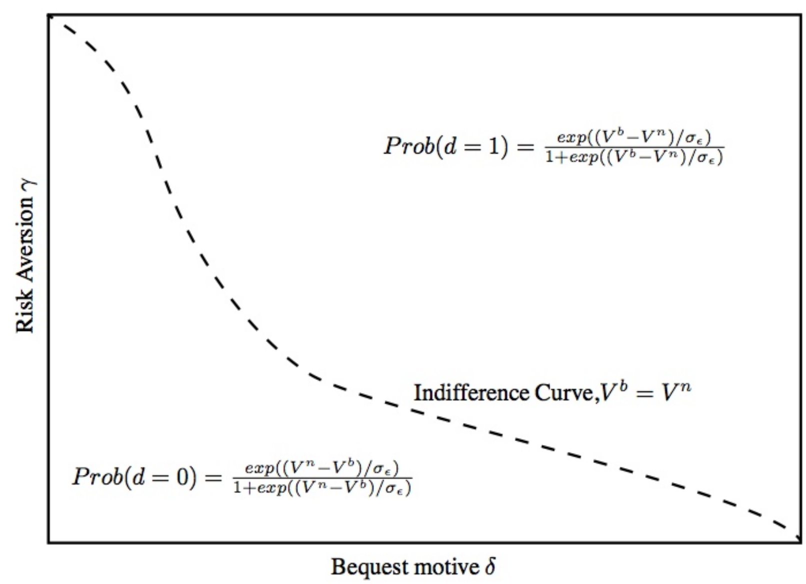


Figure 4.2: Optimal Decision With Utility Shocks



Assumption. *The purchasing decision d_i is optimally chosen by individuals described in the model.*

Assumption. *Preference distribution $log(\gamma_i, \delta_i)$ is i.i.d. drawn from an unknown normal distribution $N(\mu, \Sigma)$, where $\mu_i = (\mu_\gamma, \mu_\delta)$ and $\Sigma = \begin{bmatrix} \sigma_\gamma^2 & \sigma_{\gamma\delta} \\ \sigma_{\gamma\delta} & \sigma_\delta^2 \end{bmatrix}$ is the variance and covariance matrix of (γ, δ).*

Assumption. *The premium $p(a)$ and nursing home care expense B satisfy the condition: for all the a, N and w,*

$$\begin{aligned} &V_2'(a, w + (1 + r) * p(a - 2), h, D = 1, N, M = 0) \\ &\quad \leq V_2'(a, w, h, D = 0, N, M = 0) \end{aligned} \tag{4.15}$$

where V_2' is the first order derivative of V w.r.t. the second term w.

The first two assumptions are standard. Without assumption 1, the data can not reveal any useful information about preference under the model setup. Without the distribution assumption, we could not evaluate the likelihood of the observed data. Note assumption 2 also implies $\gamma_i > 0$, that is seniors are risk averse. The crucial assumption is the last one, which literally states the benefit of getting covered currently by paying the premium in the last period is big enough in the sense that the marginal value of liquid assets under coverage, even compensated by the premium paid $p(a - 2)$) is less than the marginal value of liquid assets if not covered by any insurance. In a very coarse understanding, the premium $p(a)$ and the nursing home care expense B under current Medicaid rules should make getting covered attractive marginally. When $a = 2$, the condition holds trivially given $b()$ is concave. To fully appreciate assumption 3 is difficult without presenting the key proposition for the identification. Thus, we present the key proposition and then move back to discuss the last assumption.

Lemma. *If the above three assumptions hold, then $d_i = 1$ if and only if (δ_i, γ_i) is above the indifference curve (δ^*, γ^*) , which is the locus such that $V_i^b(a, w, h, D, N = 0, M = 0) = V_i^n(a, w, h, D, N = 0, M = 0)$ for* *all $a \leq A - 1$.*

Proof. The difficulty of the proof lies in the fact that the value function doesn't have a closed form solution[10]. To overcome it, we exploit the finite horizon recursive structure of the model. We omit the state variables M and N since both take the value zero. Consider the case when $D = 0$. Let $V^* = V^b(a, w, h, D = 0) = V^n(a, w, h, D = 0)$ at some value γ^*. The FOC of V^b gives $u'(c^b) = \beta(1 + r)EV_2'(a', w', h', D = 1, N')$ and FOC of V^ngives $u'(c^n) = \beta(1 + r)EV_2'(a', w', h', D = 0, N')$. Consider three cases. First, $c^b = c^n$, which can't be true by assumption 3. Second, let $c^b < c^n$. It implies $u'(c^b) > u'(c^n)$ and $w'^b > w'^n - p(a)(1 + r)$, again this is impossible by assumption 3. So, the only case is $c^b > c^n$ at γ^*. Now consider a very tiny increase in γ^*, say $d\gamma$. Since it is very tiny, the increase only has first order effects on the value of $u()$and $V()$, and negligible second order effect on $c^b$and c^n. Since $u'_\gamma < 0$, a $d\gamma$ increase will drop the value of $u(c^b)$ and$u(c^n)$. And since $u''_{\gamma c} < 0$, $u(c^b)$ drops less than $u(c^n)$. Since the value function $V()$ inherits the properties of the utility function $u()$, a similar result can be found for the second term of the $V^b()$ and $V^n()$. Combining these two gives that $V^b() > V^n()$ after a tiny increase in γ^*. For values $\gamma > \gamma^*$, the same conclusion can be made based on the same reasoning. A similar result can be found for $D = 1$. Finally, we have not yet shown the existence of γ^*. However, this can be easily achieved by noticing the monotonicity of the value function, which is omitted. □

From the above proof, it is clear that assumption 3 is critical in building the conclusion that $c^b > c^n$ on the indifference curve. Notice that c^b is the optimal consumption level if a senior chooses to buy an insurance policy, while c^n is the optimal consumption if the senior does not buy. In a static world, the optimal consumption of not buying should be larger than the optimal consumption of buying since the latter has to pay an extra premium. In a dynamic world, it does not hold anymore because to make a buyer and

[10] A strict proof is very difficult to retain therefore the provided proof is very heuristic. However, the property is validated during empirical analysis.

a non-buyer indifferent on lifetime utility, the non-buyer has to save more to compensate for the utility at all future uncertain life periods.

The above proposition presents a monotonic decision rule that jumps at some indifference curve (γ^*, δ^*). However, it literally implies there are no unobservable utility shocks and decisions strictly follow a deterministic rule. In reality, a deterministic decision model fits data less successful than a model allowing utility value shocks. Following the work of Rust [1987], we assume the following decision rule

$$d = 1(V^b + \epsilon^b - (V^n + \epsilon^n)0) = 1(\epsilon^n - \epsilon^b \leq V^b - V^n)$$

where ϵ^b and ϵ^n are shocks to utility values V^b and V^n and follow type I extreme value distribution with variance $\sigma_\epsilon^2 = \frac{\pi^2}{6}$. Thus, $\epsilon = \frac{\epsilon^b - \epsilon^n}{\sigma_\epsilon}$ follows the logistic distribution, which gives the probability of choosing $d = 1$,

$$Prob(d = 1) = \frac{exp((V^b - V^n)/\sigma_\epsilon)}{1 + exp((V^b - V^n)/\sigma_\epsilon)}$$

Figure 4.2 shows the random decision rule where the probability of buying a policy at each point is given by a logit model.

4.4.3 Belief Estimate and Parameters Calibration

Accurately pinning down beliefs on risks of entering into nursing home facility, mortality and long term care expense is necessary for identifying preference parameters. Generally, belief is a result of information retrieving and evaluating process with abundant unobservables. As discussed in chapter 3, this complicated process is usually simplified to an *ex post* calibration under rational expectations assumption, which essentially states that observed *ex post* occurrence is good enough to infer *ex ante* expectations. This is usually the only practical choice since in most situations no subjective data with respect to individual beliefs is available. However, one feature of the HRS data is the extensive, detailed information on seniors' subjective probabilities on various uncertainties like long term care and life expectancy.

Belief on nursing home expense B is not asked in the HRS data set.

To overcome the problem, we assume that belief is rational: the expected expense B is identical with the true expenses for individuals ever admitted into a nursing home facility. Since the true expense is not known for individuals never admitted into a nursing home facility, we impute the expense B by adopting a two-step method.

Why Subjective Risks The objective of the paper is to understand how seniors make long term care insurance decisions by recovering risk and bequest preferences from observing insurance purchasing choices made by forward-looking rational seniors. As Manski [2004] argued the classical identification strategy hinges on the rational expectations assumption, which assumes *ex ante* belief identical with risk inferred from the *ex post* realization (rational risk). However, the previous chapter clearly shows that seniors systematically misunderstand some information and form a biased *ex ante* subjective expectation about future nursing home utilization. Spinnewijin [2009] found evidence in support of the biased perceptions of risk by showing empirically that unemployed workers overestimate how quickly they will find work, but underestimate the return to search efforts. Snowberg and Wolfers [2010] found evidence in favor of the view that mis-perceptions of the winning probability drive the favorite-long shot bias in racetrack gambling markets. The discrepancy between the *ex ante* subjective belief and rational risk presents a serious challenge to the traditional treatment.

Besides the expectation bias, beliefs inferred from the rational expectations assumptions does not capture private information an individual holds while forming his/her subjective beliefs. For example, it is very common in our data set to observe two individuals with identical observable characteristics reporting different subjective risks about future nursing home need and mortality. Finkelstein and McGarry [2007] compared individuals' subjective risk of entering a nursing home with the risk assessed by insurance companies' risk table and found that individuals have some residual private information regarding the risk. With the presence of expectation bias, it is conceptually inferior to calibrate the decision process without capturing private information.

Last, even if *ex ante* subjective risk is free of biases, using *ex post* realization to assist identification is inappropriate because of moral hazard. If seniors under insurance coverage are more prone to enter into a nursing home, belief inferred from *ex post* realization, no matter what expectation assumptions were made, is always inconsistent because the realization includes the moral hazard information as well. More specifically, consider an example in our context where a senior with rational expectations and moral hazard makes a decision to buy a LTC insurance policy. The realization of the random event is more likely to be observed because of moral hazard. Without taking it into account, the inferred *ex ante* risk will be higher than and thus inconsistent with the true *ex ante* belief on which the senior's decision is based [11]. Since moral hazard is a rampant phenomena in various insurance markets, a traditional treatment of assuming its absence would not obtain consistency.

The major challenge to subjective risk is the validity of subjective data. However, chapter 3 has shown subjective risk actually predicts realization pretty well, and evidence of good or better performance of self reported subjective probability in explaining observed choices has been confirmed in other research. For example, Juster [1966] found self reported purchasing probability actually outperforms the verbally assessed buying intentions in explaining the automobile purchase behavior. Hurd and McGarry [1995] found the HRS respondents self reported subjective probability of survival aggregate to the actual population probability and predict the actual survival. Hurd and McGarry [2002] presented evidence that the respondents actually modify their survival probability in responding to new information and the subjective survival probability do predict actual survival: those survived in the survey panel reported about 50% greater probability than those died. In experiments, Nyarko and Schotter [2002] and Bellemare, Kröger and Van Soest [2008] found models using subjective probability data can generate much better sample prediction than various "rational" models where *ex*

[11] I take a naive understanding about rational expectations here. A sophisticated rational individual might also take the moral hazard into account while forming the rational expectation.

ante probability is inferred from the *ex post* realizations.

To summarize, we argue that subjective risks are a better choice than risks inferred from the rational expectations assumption in recovering primitive parameters and therefore both subjective nursing home risk and subjective mortality risk will be utilized in all our analyses.

Transforming Subjective Risks Though self reported subjective risks utilize all private information, they are not designed for our analysis. Specially, in the HRS data set, the subjective risk of entering a nursing home facility is framed in a five year horizon while the mortality risk is in different horizons that depend on ages of seniors. In contrast, a period in our model corresponds to two years. Therefore, subjective probabilities defined in other horizons need to be transformed to probabilities defined over a two year horizon. Typically, the subjective belief of nursing home care is elicited by a formal specification of risk range from 0% to 100% and the following survey question:

> What is the chance that you will move to a nursing home in the next five years?

which is immediately followed with a definition of nursing home[12]. Similarly, the subjective belief of mortality is asked by the following survey question:

> What is the percent chance that you will live to be 75 or more?

Obviously, the subjective belief on nursing home care is asked in a five year horizon while the mortality belief is surveyed in various horizons depending on a senior's age at the survey. In the following discussion, we show how to transform the self-reported belief defined in a five-year horizon into a two-year belief without losing the embedded private information. In our view, there are two minimum standards that must be met for any reasonable transformation: first, the resulted probabilities should have the range

[12]Chapter 3 gives details on subjective question on nursing home.

from 0% to 100%. Second, for any two probabilities p_1 and p_2, the after-transformation probability of the former should be no less than that of the second one, $tr(p_1) > tr(p_2)$, if $p_1 > p_2$. In other words, the transformation should preserve the order of probabilities.

Our starting point is to assume a hazard analysis process while an individual assessing probabilities. Borrowing the jargon of survival analysis, an individual reports a risk by integrating hazard rate given by the mixed proportional hazard (MPH) model :

$$h_i(t) = a_i h_0(t) exp(\beta_0 + X_i \beta_1)$$

where $h_i(t)$ is the individual i's hazard rate at time t, a_i is idiosyncratic hazard factor and represents the private information the individual holds, $h_0(t)$ is the common hazard rate for all individuals in the economy and X_i is the individual's observable characteristics. To see how the MPH model can help in preserving the private information, notice the subjective risk is actually the cumulative probability$F_i(t)$, which is connected to the hazard rate though

$$-ln(1 - F_i(t)) = \int_0^t h_i(v)dv = \int_0^t a_i h_0(v) exp(\beta_0 + X_i \beta_1)dv$$

The primary objective here is to infer $F_i(2)$ from $F_i(5)$ since a period in our model corresponds to two years. To recover $F_i(2)$, we assume $h_0(t) = h_0$, which implies a constant baseline hazard factor. This seems, in our setup, an acceptable assumption. It states that the common hazard rate of the whole society is independent of the duration. Under the constant common hazard rate assumption, the risk is given by

$$-ln(1 - F_i(t)) = a_i exp(\beta_0 + X_i \beta_1) h_0 t$$

and $F_i(2)$ is thus given by

$$F_i(2) = 1 - (1 - F_i(5))^{2/5}$$

and similarly for other horizons.

The proposed transforming method meets the standards discussed above. First, it preserves the fundamental property of probability: the transformed probability has a range from zero to one. Second, it also preserves the private information an individual holds in assessing the subjective probability in the sense of maintaining the order of probabilities. An other advantage is that it exploits the key idea of survival analysis and does not impose strong assumptions.

Last, we maintain the independent assumption of risk events. Allowing dependence between nursing home risk and mortality risk requires to specify how subjective beliefs interact. This is not an easy task without assuming how seniors form and update subjective assessments.

Though subjective risks of both nursing home care and mortality are observed for each individual at current age, the beliefs for the rest of life are not available. In the simplest case, we could assume identical belief of nursing home care or mortality at later life stages. Note, *belief for the rest of life* is what individuals hold at current age for later life stages, not subjective belief held at later life stages. For nursing home risk, it seems acceptable to assume it does not vary with age. However, for mortality risk, it ultimately says individuals believe the risk of death is independent of age.

Intuitively, there are two methods available to better recover the later subjective mortality beliefs. The simplest one is to apply a nonparametric matching process, where for each individual, according to his/her other attributes like gender, education and income *etc.*, his/her subjective mortality beliefs at a later stage of the life a' are the mean value of the reported subjective mortality risks reported by those who are aged at a'. However, this method ignores life expectation improvement for younger cohorts and more importantly, ignores the private information of the individual included in his/her current subjective risks. Another method, which is adopted in this chapter, to remedy this problem is to recover idiosyncratic mortality risks at later life stages from the reported subjective risk of mortality at current age by assuming an evenly increasing rate until the maximum age A at which the subjective probability should be one. Other methods of assuming an accelerating increase or utilize other demographics characteristics are also considered in our computation.

Imputation of the Belief of Expense B Since the expense B belief is not asked in the HRS data set, we have to impute it somehow. For those ever admitted into nursing home care during our study period, the expense B belief is assumed identical with true expense. For individuals never admitted, the expense B belief is imputed using outside data sources [13]. During the period of writing the thesis, the only available data set augmented with detailed health information and long term care expenditure is the National Long Term Care Survey (NLTCS) 2004. It is a longitudinal survey conducted by Duke University and designed to study changes in health and functional status of old Americans.

In the first step, we use the NLTCS 2004 data to estimate the prediction equation for nursing home care expense. The equation is given by

$$B = \beta_0 + X'\beta_1$$

where X is a vector of personal observables, which consist of six health measurements and four demographic variables. The vector X is observed in both data sets so that we can use the estimated prediction equation from the above NLTCS data to impute the $E(\hat{B})$ for each individual in the HRS sample. At the second step, the expense B belief is assumed identical with the estimated value $E(\hat{B})$. Equalizing the projected expense and the belief seems restrictive because it implies that people form expectations following the OLS regression. However, the previous chapter has argued under rational expectations, the coefficient of X should be identical in both belief and true expense regressions. Hurd and McGarry [1995] noticed that the morality beliefs in HRS actually covary with other variables in the same way actual outcomes vary with the variables in OLS regression. Further, the proposed method is only one convenient way for the purpose of identi-

[13] Of course, we prefer to use the realized expense inside the HRS data set to project the expense. However, the admitted sample is too small to give a reasonable estimate.

fication: any other methods that help to infer beliefs B would work as well. As a robustness check, we examine an alternative method of nonparametric estimate of B for all seniors based on age, gender and education and do not find significant effects on our following main results. Finally, for the expected nursing home expenses at a later life stage, we assume each individual imposes a yearly 3% growth rate.

Calibration of β, r , $\underline{c}$, A and $p(a)$ We treat the parameters β and r as observables and calibrate their values. The discount factor β is assumed to be fixed at 0.98, and r is set at a value such that $\beta * r = 1$. The maximum age A is set to be 100. Finally, we set $\underline{c}$ at 1000 dollars a month, approximately the Department of Human and Health Service poverty guideline level for one person household.

Since premiums are only known for these who bought an LTCI, we have to infer potential premiums for these who did not buy any private policies. To do this, we use a nonparametric matching method by noticing that premiums only depend on age and gender in each local market. We therefore match the premium for each individual to the estimated mean premium for each subgroup determined by the three factors.

4.4.4 Likelihood Function

The likelihood function for each observation is given by

$$\begin{aligned}
&l_i(d_i; \mu, \Sigma) = \\
&1(d_i = 1) * log[\int_{\infty}^{\infty}\int_{\infty}^{\infty} \frac{exp(V^{b-n}(\gamma, \delta))}{1 + exp(V^{b-n}(\gamma, \delta))} \phi(\gamma, \delta) d\gamma d\delta] + \\
&1(d_i = 0) * log[\int_{\infty}^{\infty}\int_{\infty}^{\infty} \frac{1}{1 + exp(V^{b-n}(\gamma, \delta))} \phi(\gamma, \delta) d\gamma d\delta]
\end{aligned} \tag{4.16}$$

where 1() is an index function, $V^{b-n}(\gamma, \delta) = \frac{V^b(\gamma,\delta) - V^n(\gamma,\delta)}{\sigma_\epsilon}$ is the Bellman value difference between buying and not buying decisions at each preference set (γ, δ) divided by the scale factor σ_ϵ and $\phi(\gamma, \delta)$ is the probability density function of the bivariate normal distribution of $log(\gamma, \delta)$. Thus, we have 6 parameters to estimate: scale factor σ_ϵ, mean μ_γ and μ_δ, variance σ_γ and σ_δ, and covariance $\sigma_{\gamma\delta}$. Ideally, we would prefer to have a closed form, continuous function of $V^b(\gamma, \delta) - V^n(\gamma, \delta)$ for the purpose of integration. However, it is not possibly achieved because of the unavailability of the closed form solution to the Bellman value functions of V^b and V^n. Thus, we choose to calculate the likelihood by dividing the γ and δ space into many small grids and calculating the cumulated probability of the small grid. Since the grid is very small, we thus can approximate the Bellman value differences in the grid by the crossing middle point of the grid. In this way, the likelihood is calculated as

$$\begin{aligned} & l(d_i; \mu, \Sigma) = \\ & 1(d_i = 1) * log[\sum_{j=1}^{n_\gamma} \sum_{k=1}^{n_\delta} \frac{exp(V^{b-n}(\gamma_j, \delta_k))}{1 + exp(V^{b-n}(\gamma_j, \delta_k))} \Phi_\Delta(\gamma_j, \delta_k)] + \\ & 1(d_i = 0) * log[\sum_{j=1}^{n_\gamma} \sum_{k=1}^{n_\delta} \frac{1}{1 + exp(V^{b-n}(\gamma_j, \delta_k))} \Phi_\Delta(\gamma_j, \delta_k)] \end{aligned} \tag{4.17}$$

where $V^{b-n}(\gamma_j, \delta_k)$ is the Bellman value difference at the grid (j, k) and $\Phi_\Delta(\gamma_j, \delta_k)$ is the cumulative probability of the grid [14]. When the grid size goes smaller, we can approximate the likelihood value closer but with more computational burden.

To numerically calculate the likelihood, we first need to define an appropriate range of γ and δ. The choice of bound limit of γ and δ is through a guess and verification process: pick some reasonable range and do the maximum likelihood estimation. If the estimated mean and standard error shows the cumulative probability of the square defined by the range is at least 0.99, then the range is chosen as the satisfied one; Otherwise, we increase the bound of the grid square. After defining the range of γ and δ, we need to select the appropriate step size to define the grid. At the initial attempt, we choose a random step-size, say 0.5 and do the estimation, if

[14] Or precisely, $\Phi_\Delta(\gamma_j, \delta_k) = \Phi(\gamma_{j+1}, \delta_{k+1}) - \Phi(\gamma_j, \delta_k)$, the difference of CDF at the grid.

the estimated probability at each grid is less then 0.01, then the selected step-size is the appropriate one; otherwise, we refine the step-size. While this increases the computation, the selection can give us more precise measurement on the likelihood. With all these considerations, the range of $[\underline{\gamma}, \overline{\gamma}]$ and $[\underline{\delta}, \overline{\delta}]$ is set at $[-5, 5]$ for $log(\gamma)$ and $[-10, 10]$ for $log(\delta)$ and the step size is 0.1. Thus, the grid is a 100 by 200 matrix for each individual. Notice, $exp(-5) = 0.0067$ and $exp(5) = 148$, which should include all reported risk aversion level in literature.

4.5 Results

4.5.1 The Distribution of δ and γ

Table 4.3 reports the estimated distribution of the preferences and the variance of utility shocks, σ_ϵ. The estimated $\mu_{ln(\gamma)}$ is 0.2786, which corresponds to a mean value of γ 1.322. It is close to other estimated risk aversion using different methods and data sets. For example, Chetty [2006] reported an upper bound of $\gamma < 2$ on the curvature of utility over wealth by exploiting the link between risk aversion and labour supply behaviour. The variance of $ln(\gamma)$, $\mu_{ln(\gamma)}$, is only 0.0535. The small variance possibly reflects the sample selection in our analysis: the seniors are singles with substantial assets and similar educations. The bequest motive after taking the logarithm is estimated at -3.3801, which is equivalent to $\delta = 0.0341$. This value is far less than 1. Since in our setup, the bequest motive is similar to a scale factor to the utility function, the small estimated δ implies that individuals treat unconsumed bequests as an *inferior* form of consumption. The small magnitude of the bequest motive seems to support the accidental bequest theory rather than the altruistic model. If bequest is driven by altruism to offspring, it seems reasonable to expect it to be close to 1: the person making the bequest prefers consumption left to his or her offspring at least the same as his or her own consumption. Indeed, Hurd [1989] found a similar conclusion by estimating a life cycle model of consumption. He found the marginal utility of bequest is small, therefore he suspected most bequests

Table 4.3: Preference Distribution Estimates

	Estimated Valued	Standard Error
$\mu_{ln(\gamma)}$	0.2786	0.0223
$\mu_{ln(\delta)}$	-3.3801	0.8976
$\sigma_{ln(\gamma)}$	0.0535	0.0029
$\sigma_{ln(\delta)}$	0.8314	0.5121
ρ	0.0801	0.0373
σ_{ϵ}	0.0116	0.0103
Obs	696	696

are accidental. The variance of $ln(\delta)$ is 0.8314. Combined with the mean value $\mu_{ln(\delta)}$, it implies a upper tail value of $ln(\delta)$ at 95% is merely -1.72, corresponding a value of δ at 0.18. It is still far less than 1. The magnitude of $\sigma_{ln(\delta)}$ is more than 10 times of that of $\sigma_{ln(\gamma)}$. Thus, we estimate substantial heterogeneity across individuals in their bequest motives. The correlation ρ is 0.0801. Seniors have stronger bequest motive are also tend to be more risk averse, however the correlation is very weak.

4.5.2 Evaluating *ex post* Preferences

The *ex post* Expected Risk Aversion γ and Bequest Motive δ One way to evaluate the model and estimated preference distribution is to calculate the *ex post* expected risk aversion and bequest motives for each individual and compare them with other sources in the data. In the HRS data set, seniors were asked to do some hypothetical job games which can be used to elicit risk aversion. Further, information on children and subjective probability to leave a bequest offers a good opportunity to evaluate the estimated *ex post* bequest motive.

The *ex post* expected risk aversion is defined as

$$E(\gamma|d:\mu,\Sigma) = \sum_{j=1}^{n_\gamma} \gamma_j p(\gamma_j|d:\mu,\Sigma)$$

where $p(\gamma_j|d:\mu,\Sigma)$ is the conditional probability of γ fall within grid j

given values of d, μ, and Σ, which is defined as

$$p(\gamma_j|d:\mu,\Sigma)=\sum_{k=1}^{n_\delta} p(\gamma_j|d,\delta_k:\mu,\Sigma)p_\delta(\delta_k:\mu,\Sigma)$$

where $p_\delta(\delta_k|\mu,\Sigma)$ is the marginal probability of δ within grid δ_k. Notice $p(\gamma_j|d,\mu,\Sigma,\delta_k)$ is by Bayes rule

$$p(\gamma_j|d,\delta_k:\mu,\Sigma)=\frac{p(d|\gamma_j,\delta_k:\mu,\Sigma)p(\gamma_j,\delta_k:\mu,\Sigma)}{p(d,\delta_k:\mu,\Sigma)}$$

Notice $p(d,\delta_k:\mu,\Sigma)$ is given by

$$p(d,\delta_k:\mu,\Sigma)=p(d|\delta_k:\mu,\Sigma)p_\delta(\delta_k:\mu,\Sigma)$$

and $p(d|\delta_k:\mu,\Sigma)$ can be expressed as

$$p(d|\delta_k:\mu,\Sigma)=\sum_{j=1}^{n_\gamma} p(d|\gamma_j,\delta_k:\mu,\Sigma)p_\gamma(\gamma_j:\mu,\Sigma)$$

where $p_\gamma(\gamma_j:\mu,\Sigma)$ is the marginal probability of γ within grid γ_j. Thus, the *ex post* expected γ is

$$E(\gamma|d:\mu,\Sigma)=\sum_{j=1}^{n_\gamma}\gamma_j\sum_{k=1}^{n_\delta}\frac{p(d|\gamma_j,\delta_k:\mu,\Sigma)p(\gamma_j,\delta_k:\mu,\Sigma)}{\sum_{j=1}^{n_\gamma}p(d|\gamma_j,\delta_k:\mu,\Sigma)p_\gamma(\gamma_j:\mu,\Sigma)}$$

The *ex post* expected bequest motive is defined similarly.

Hypothetical Job Game and *ex post* Risk Aversion γ In the HRS data set, seniors were also asked to make a choice between a certain job and a risk job under hypothetical scenarios. The hypothetical game was introduced by the following words:

> Now I have another kind of question. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current

> total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take – the first job or the second job?

If the respondent accepts the risky job, she/he is proposed to choose between:

> Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?

If the choice is still the risky job, the respondent is asked about the riskiest job:

> Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it by seventy-five percent. Would you take the first job or the second job?

Similarly, if the respondent's choice to the initial question is the certain job, she/he is asked a less risky alternative:

> Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?

and the least risky alternative:

> Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it by ten percent. Would you take the first job or the second job?

Based on the answers to the job game questions, we can calculate the *ex post* risk aversion grouped by the answers and test if seniors choosing less risky options have higher *ex post* risk aversion. Since very few subjects choose other options but the least risky alternative (277 among 696), we group all subjects choose the three riskier options together. Among these choosing the least risky option, the average risk aversion γ is 1.23 with a standard error of 0.49. The average risk aversion among those choosing riskier options is 1.19 with a standard error of 0.45. Our estimates pass the test that those choosing the least risky option have a higher average risk aversion.

However, our *ex post* risk aversion is much smaller than that implied in these hypothetical games. By assuming a CRRA utility function over wealth $u(w) = \frac{w^{1-\gamma_i}-1}{1-\gamma_i}$, it is possible to recover the risk aversion range for each respondent. For example, for a respondent rejecting a one-third downside job but accepting a one-fifth downside job, the upper and lower bound the respondent's risk aversion are

$$0.5 * \frac{2^{1-\gamma}-1}{1-\gamma} + 0.5 * \frac{(2/3)^{1-\gamma}-1}{1-\gamma} = \frac{1^{1-\gamma}-1}{1-\gamma} \Longrightarrow \underline{\gamma} = 2$$

$$0.5 * \frac{2^{1-\gamma}-1}{1-\gamma} + 0.5 * \frac{(4/5)^{1-\gamma}-1}{1-\gamma} = \frac{1^{1-\gamma}-1}{1-\gamma} \Longrightarrow \overline{\gamma} = 3.76$$

Since most the of respondents choose the least risky option, we thus expect an average risk aversion greater than 3.76, which is much bigger than the estimated mean of the risk aversion distribution. In a recent paper, Einav and Finkelstein *et al.* examined individual's choices over several employer-provided insurance coverage options and one 401(k) investment. Essentially, they found individuals' choices in one of the insurance domains are more predictive of the other insurance policies than of the 401(k) investment decision. Our results, however, seem to support that risk aversion estimated from different domains actually has the same order but with different magnitudes.

Children, Subjective Bequest-Leaving Probability and *ex post* Bequest Motive δ It is popularly believed that bequest motive should correlate with the presence of children. Under the altruistic theory of bequests, it is natural to predict this correlation. On the other hand, if the driving factor of bequests is for consumption smoothing and bequest motive is purely accidental, it is not clear why the presence of children should be a good predictor of bequest motive. Empirically, Hurd [1987] analyzed the bequest motive by using a ten year panel and found no wealth changes differences between households with living children and without living children. Similarly, Jurges [2001] found that having children has no significant impact on households' wealth trajectories in German. We have a similar finding here. By grouping the seniors over the presence of living children, we ended with 89 seniors without any children and 586 seniors with at least one children. The average bequest motive for these with at least one child is 0.052(0.049), which is very close to the 0.051(0.050) of these without any children [15].

This finding arises an interesting question that is why people appreciate money left after death if there is no offspring to take the bequest [16]. Of course, a reasonable justification would be the presence of nieces or nephews that would potentially substitute offspring. More broadly, charity donations, psychological factors and habitual preference would also be the potential reasons. To further explore whether the presence of offspring have substantial effect on estimated bequest motive, we reestimate the preference parameters based on the presence of children. The estimated bequest motive $\mu_{ln(\delta)}$ is $-3.8543(1.2003)$ for those without any children *vs.* $-3.1003(0.9327)$ for those with at least one child, and the estimated risk aversion $\mu_{ln(\gamma)}$ is 0.2865(0.0402) for those without any children *vs.* 0.2673(0.0314) for those with at least one child. Though the closeness of risk aversion is of expected, the independence of bequest motive and children is not surprising: it further confirms the previous accidental bequest motive conclusion.

We last test if the *ex post* bequest motive can be predictive of various self reported bequest leaving probabilities. In the HRS data set, seniors were asked to report the probability of leaving any bequest, the probability

[15] Another related concern is that these seniors with children living close by can have an alternative to home care given by children. A similar analysis shows that the estimated bequest motive for these with at least one child living close by is 0.056(0.054), which is very close to 0.051(0.057), the distribution for these without any children living close by.

[16] Notice the specification of the bequest utility function 4.14 does not depend on the presence of children and the bequest motive is defined as a scale parameter to the usual CRRA utility function.

Table 4.4: How Bequest Motive Predicts Subjective Bequest Probability?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Bequest Motive δ	34.96***	38.46***	23.77**	53.53***	56.60***	33.46*	105.68***	113.93***	63.63**
	(10.81)	(11.71)	(11.32)	(17.33)	(17.97)	(18.35)	(29.23)	(37.50)	(28.19)
Controls:									
Age and Gender		YES	YES		YES	YES		YES	YES
Asset, Income and Health			YES			YES			YES
R^2	0.01	0.02	0.06	0.01	0.02	0.07	0.02	0.04	0.25
Observations	418	418	418	680	680	680	671	671	671

Notes: Standard errors are in parentheses. Column $1-3$ corresponds to leave any bequest, $4-6$ to leave bequest at least $\$10K$, $7-9$ to leave bequest at least $\$100K$. Dependent variable is subjective bequest probability. Bequest motive is the *ex post* estimated value. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

of leaving bequest greater than $\$10k$ and the probability of leaving bequest greater than $\$100k$. Though the mechanism and underlying factors influencing subjective bequest leaving probability are unknown, the subjective probability of leaving a bequest should be a function of bequest motive, household assets and income *etc.*. Table 4.4 presents the regression results of various bequest leaving probabilities on *ex post* bequest motive. For example, column (1) of the table lists a coefficient of 34.96, which means an increase in posterior expected bequest motive by 1 unit corresponds to an increase of the probability of leaving any bequest by 34.96 percent. The expected posterior bequest motive is 0.051 with the same size standard deviation. It implies a senior with one standard deviation increase in posterior expected bequest motive generally reports a higher probability of leaving any bequest by about 1.75%. Similar reading can applied to other columns. Obviously, in all of the regressions, the *ex post* bequest motive is a significant predictor of various bequest leaving probabilities.

4.5.3 Understanding the Puzzle

One motivation of this chapter is to understand the long term care insurance market puzzle where a positive correlation between LTC insurance coverage and LTC incidence rate is not supported by data. One explanation is there exists a second dimension of private information that negatively correlates with risk status and therefore offsets the positive correlation. And our estimate shows a significant heterogeneity of the bequest motive other than risk aversion among seniors.

To formally compare the two factors in explaining the absence of positive correlation, we employ a general approach proposed by Chiappori and Salanie [2000]. This approach estimates a bivariate probit of insurance coverage and risk incidence conditional on all variables X observed by insurers. Let y and z be dummy variables of insurance coverage and risk occurrence respectively and ψ and η be a pair of jointly normal distributed errors, then

$$\begin{cases} y = \mathbf{1}(X\beta + \psi > 0) \\ z = \mathbf{1}(X\beta + \eta > 0) \end{cases}$$

where $\mathbf{1}(.)$ is an index function. Conditional on X observed by insurers, the classical asymmetric information theory predicts a positive correlation between ψ and η because of either adverse selection or moral hazard. It can be shown the estimated coefficient between ψ and η follows a χ^2 distribution with degree 1. The empirical lack of positive correlation leads to the suggestion that private information on preferences, in addition to private information on risk status, is another major driving factor in LTCI shopping decisions. If so, a positive correlation should be observed after controlling for the factor in the insurance shopping equation y (whether the preference should be controlled in the risk incidence equation is irrelevant since it should not affect risk incidence).

Following Finkelstein and McGarry [2006], control variable X includes observable individual health and demographic status and risk classification assigned by insurance companies. Table 4.5 presents the estimated coefficient and Wald test value of the coefficient is zero. The first column confirms the lack of positive correlation in the LTC insurance market if no preference is controlled. In the second column, we further control for the *ex post* risk aversion. The coefficient is still small and the χ^2 value is too small to reject

Table 4.5: Positive Correlation Test

	(1)	(2)	(3)	(4)
Correlation between ψ and η	.071	.058	.513	.460
Wald Test of the correlation=0 ($\chi^2(1)$)	.374	.233	3.852**	3.484*
Controls:				
Insurer Observables	YES	YES	YES	YES
Risk Aversion		YES		YES
Bequest Motive			YES	YES
Observations	660	579	660	579
Wald χ^2 value	3762.13	3490.11	7018.31	5583.58

Notes: This table lists the estimate results from biprobit regressions. The dependent variables are LTCI coverage and LTC risk incidence. Insurer observables include 21 demographic and healthy conditions. A single asterisk denotes significance at the 10% level, double for 5%, and triple for 1%.

the zero hypothesis. The third column controls for the insurer observables plus an *ex post* bequest motive. The coefficient jumps to 0.513 with a χ^2 indicating a significance level of 5%. The last column further controls for risk aversion and the last result maintains. Consistent with our previous finding, the results indicate that the bequest motive actually has a more substantial impact on LTCI shopping decisions.

This finding clearly shows the heterogeneity on bequest motive is an important factor in formulating seniors' insurance shopping decisions. Combined with the small mean value of the bequest motive, we suspect the decisions are mainly driven by a strong desire to smooth consumption rather than a altruistic one to leave offspring assets. The heterogeneity on risk aversion is less significant quantitatively and less important in explaining the puzzle.

4.5.4 Counterfactual Policy Analysis

In spite of high long term care cost, only about 10 percent of the seniors are covered by private insurance policies. Even among the highly selected sample in our analysis, only 17% bought any policies. This thin market has been extensively discussed in literature. Pauly [1990] argued that a LTCI coverage's primary object is to protect bequest, but it is already excessive because of imperfect annuities. Further, the substitution of formal care to children care makes a policy also less favorable. Brown and Finkelstein [2008] examined the interaction of the public Medicaid program with the private market for long-term care insurance and found that Medicaid could explain the lack of private insurance purchases for about two-thirds of the wealth distribution even if there were no other factors limiting the size of the market. Goda [2010] exploited variation in adoption and generosity of state tax subsidies for private long-term care insurance and found the average tax subsidy raises coverage rates 28 percent.

In this section, we do two counterfactual policy analyses to shed light on the thin market size issue. First, we consider a policy that subsidizes insurance premiums by taxing bequest. During the study period, the federal bequest tax policy in the United States only taxes estates valued greater than $\$675,000$ with many exemptions. Taxing all estates literally increases the price of bequests and makes consumption and LTCI less expensive. However, it is possible for the income effect to dominate the substitution effect. The subsidy to LTC insurance premium makes purchasing coverage more attractive. Second, we can alternatively subsidize the LTC cost instead of the premium by taxing bequests. The insurance coverage becomes less attractive if the cost of LTC is smaller. On the other hand, income effects will increase purchasing power and thus the market size.

To conduct the analysis, we recalculate the probability of buying an insurance policy for each individual in order to get an average probability. We do this both under the cases without any policy interventions and these with the proposed policy interventions. The policy effect is defined as the ratio of the difference between average probabilities to the probability without policy interventions. Table 4.6 presents the policy effects under various policy combinations. As expected, the market size increases with the subsidy on LTCI premiums increase. Actually, the market size would increase 14.86% at the subsidy rate of 20% even with a tax rate of 5% on bequests. A tax on bequests harms the market size significantly: a tax rate moving from

Table 4.6: Policy Effects under Various Scenarios

Tax Rate on Bequest:	5%	10%	20%
Subsidy Rate on LTCI Premium:			
5%	-1.35%	-1.35%	-11.49%
10%	4.73%	1.35%	-7.43%
20%	14.86%	12.84%	2.70%
Subsidy Rate on LTC Cost:			
5%	12.84%	5.79%	-2.70%
10%	6.75%	2.02%	-5.41%
20%	-18.92%	-20.95%	-23.65%

Notes: Policy effect is the ratio of the probability difference between a scenario with proposed policy interventions and that without a policy intervention to the probability without a policy intervention.

10% to 20% on bequests will make the policy effect shrink from 12.84% to merely 2.70%. A Subsidy on LTC cost is more complicated because of the opposing direction of the income and substitution effects. When a subsidy rate is fixed at 5%, the policy effect of subsidizing LTC cost is bigger than subsidizing the premium. At a low level of subsidy rate, the income effects of the cost subsidy dominates the total effects of the premium subsidy. With the subsidy rate on cost increasing, the individuals find they are over paying for the same policy, which leads to the market shrinking.

The above analysis provides some insight regrading average counterfactual policy effects on population, however, it does not tell much about heterogeneous effect of the counterfactual policies across different groups defined by *ex post* expected values of γ and δ. Here, we classify seniors into four groups according to the *ex post* expected value of γ and δ using the the median values of the γ and δ as threshold values. By this way, a senior in the upper γ and upper δ group has the posterior values of γ and δ above 50 percentile. Similarly for other three groups. Table 4.7 illustrates the policy effects across four groups under various policy combinations. Rather than listing the probability difference ratio across four groups, this table reports the within-group average probability both before and after policy changes since the average level of probabilities would be very different across four groups. At the second row, it first presents the original average probability of purchasing LTCI policies across different groups. As expected, the UpperγUpperδ group has the highest level of tendency to buy a policy. Also, value of δ has a greater effect on the probability than that of γ. The average probability after a policy change is listed in following panels. Comparing the after policy average probability and the original probability across the groups, it is clear that different policies have heterogeneous effects on different groups. This is specially obvious when tax rate on bequest is 20%. With a high level tax on bequest, the incentive for the Upperδ groups to buy insurance is less strong since the motive to leave bequest is simply reduced proportionally with the tax rate. When tax rate on bequest is low at 5%, substitution effects naturally dominate income effects. This is weakly true for the policy subsidizing LTCI premium, which is found to have a greater effect on Lowerγ groups. This could be arose if lower γ values are somehow associated with the seniors who did not buy LTCI because of high premiums.

4.6 Conclusion

Understanding how seniors make long term care insurance decisions interests both policy makers and academic researchers. While long term care expense is the largest single health and financial risk facing seniors in the United States, the LTCI market is relatively small. Any public policy proposed to improve the market size and social welfare *etc.* must build on a thorough understanding of how insurance decisions are made. The lack of a positive correlation between insurance coverage and the *ex post* risk incidence rate challenges the classical prediction in the LTCI market.

This paper offers a new perspective to understand the LTCI market by constructing a structural dynamic discrete choice model. The innovation of the paper is to estimate a joint distribution of two unobservable preference parameters, risk aversion and a bequest motive. Introducing the bequest motive in the LTCI market is helpful since it has been documented exten-

Table 4.7: LTCI Purchasing Probability across Different Groups Before and After Policy Changes

Groups:	UpperγUpperδ	LowerγUpperδ	UpperγLowerδ	LowerγLowerδ
Original Probability	20.3	19.2	16.3	14.4
	Panel A: 5% Tax Rate on Bequest			
Subsidy Rate on LTCI Premium				
5%	20.0	19.1	16.2	14.4
10%	21.0	20.2	17.0	15.5
20%	22.3	21.9	18.7	17.5
Subsidy Rate on LTC Cost				
5%	22.5	21.3	18.6	16.5
10%	21.0	20.2	17.4	15.7
20%	17.3	15.9	12.8	12.1
	Panel B: 10% Tax Rate on Bequest			
Subsidy Rate on LTCI Premium				
5%	20.0	19.1	16.1	14.4
10%	20.5	19.2	16.6	14.8
20%	22.3	21.4	18.6	16.7
Subsidy Rate on LTC Cost				
5%	20.7	19.8	17.4	15.5
10%	20.5	19.6	16.8	15.0
20%	15.9	14.9	13.4	11.8
	Panel C: 20% Tax Rate on Bequest			
Subsidy Rate on LTCI Premium				
5%	17.1	16.6	15.7	13.6
10%	18.9	18.0	15.2	13.4
20%	20.2	19.3	17.2	15.3
Subsidy Rate on LTC Cost				
5%	19.2	18.3	16.4	14.3
10%	19.0	18.1	15.5	14.1
20%	14.8	14.2	13.3	11.7

Notes: This table lists the average probability of purchasing LTCI across different groups after and before a policy change defined by a combination of a tax on bequest and a subsidy on LTCI premium or on LTC cost. A group is defined by the posterior expected value of γ and δ. For example, UpperγUpperδ group consists of those with γ and δ at its higher 50 percentile. The average probability before any policy changes is listed at the second row named *Original Probability*. The average probability after a policy change is listed in Panel A, Panel B and Panel C. Therefore, 20 at column 1 and row 1 of Panel A means the average probability of purchasing LTCI is 20% for the UpperγUpperδ group after the policy combination of a 5% tax rate on bequest and a 5% subsidy on LTCI premium.

sively that a bequest motive plays an important role in explaining seniors' saving and consumption behavior.

Our estimates indicate substantial heterogeneity in bequest motive. The *ex post* estimated bequest preference is quite successful in explaining the insurance market puzzle: controlling for bequest motive generates the expected positive correlation. Thus, we conclude bequest motive heterogeneity plays a substantial role in offsetting the positive correlation. The estimate of risk aversion is consistent with other research. Because our model captures institutional details, it is capable of providing tools to evaluate potential policy effects.

Future development would be to estimate the model in a panel data framework. This requires a careful restructuring of the model for the purpose of identification and integrating fixed effects. Since our paper only focuses on a small highly selected sample, caution should be applied when associating the results to the whole senior population.

Chapter 5

Conclusion

This dissertation discusses three topics in applied economics.

The first essay examines the causal effect of social capital on individual income by exploiting the historically determined pattern of family name distribution in Chinese villages. Family name distribution impacts social capital through historical inter-lineage rivalry and cooperation. The estimates show a strong first order effect on male villagers, which implies a one standard deviation increase in social capital is equivalent to two to four years of education. No effects on female villagers were found. The gender differentiation could be accounted for by occupation difference: male villagers' income mainly comes from market exchange, while female villagers' income comes mainly from home production. Using a simple model, it is demonstrated that a village's social capital determines its trade scope and therefore income of its residents.

The second essay proposes a general method to identify subjective expectation bias. The method exploits an implication of rational expectations that requires the identical weight of an independent variable in projecting both objective and subjective probabilities. The empirical analysis shows that female seniors do not correctly internalize age information while male seniors fail at internalizing income information. Though cognitive ability and risk aversion can partially explain the results, they are not the sources of the identified biases.

The third essay explores how seniors make long term care insurance (LTCI) decisions by developing a dynamic structural discrete choice model where a rational, risk averse, bequest motivated senior has to decide at each period whether to buy an insurance policy or not [17]. Using the Health

[17]The rational model of this essay does not necessary conflict with the finding in the

and Retirement Survey data, this essay finds substantial heterogeneity in bequest motive that drives LTCI decisions. Specially, the idiosyncratic bequest motive helps to explain why LTCI holders do not experience a higher incidence rate than non-holders.

second essay. Expectation bias or rational expectation assumption are both merely one vehicle for the purpose of identification and expectation bias does not imply utility maximizing behaviour. In our view, it is very natural for people to have biased expectations and rational decisions.

Bibliography

[1] Acemoglu, Daron and Simon Johnson, 2005, Unbounding Institutions, *Journal of Political Economy*, 113, pp. 949-995

[2] Alberoni, F., 1962, Contribution to the study of subjective probability Part I. *Journal of General Psychology* 66, pp. 241-264

[3] Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R., 2003, Fractionalization, *Journal of Economic Growth*, 8(2), pp. 155-194

[4] Alesina, A., and Ferrara, E., 2002, Who Trust Others, *Journal of Public Economics*, 85, pp. 207-234

[5] Amemiya, T., 1981, Qualitative Response Models: A Survey, *Journal of Economic Literature*, Vol. 19, No. 4, pp. 1483-1536

[6] Anscombe, F. J. and Aumann R. J., 1963, A Definition of Subjective Probability, *The Annals of Mathematical Statistics*, Vol. 34, No. 1, pp. 199-205

[7] Arrow,K., 1999, Observations on Social Capital in *Social Capital: A Multifaceted Perspective* edited by Dasgupta and Serageldin, World Bank, pp. 3-5

[8] Attanasio, O. and Kaufmann, K. 2009, Educational Choices, Subjective Expectations, and Credit Constraints, NBER Working Paper No. 15087

[9] Babad, E. and Katz, Y. 1991, Wishful Thinking-Against All Odds, *Journal of Applied Social Psychology*, Vol. 21, No. 23, pp. 1921-1938.

[10] Beattie H., 1979, Land and Lineage in China: A Study of T'ung-ch'eng County, Anhwei, in the Ming and Ch'ing Dynasties, Cambridge: Cambridge University Press

[11] Bellamare, Kröger and Van Soest, 2008, Measuring Inequality Aversion in A Heterogeneous Population Using Experimental Decisions and Subjective Probabilities, *Econometrica*, Vol.76 (4), 815-39

[12] Benítez-Silva, H. and Dwyer, D. 2005, Expectation formation of older married couples and the rational expectations hypothesis, *Labour Economics*, pp. 191-218

[13] Bernheim, B. D. 1987, Social Security Benefits: An Empirical Study of Expectations and Realizations, NBER working paper 2257

[14] Bernheim, B. D. 1990, How Do the Elderly Form Expectations? An Analysis of Responses to New Information, in *Issues in the Economics of Aging*, Edited by Wise D. A., University of Chicago Press

[15] Bernheim, B.D., 1991, How Strong Are Bequest Motives? Evidence Based On Estimates of the Demand for Life Insurance and Annuities, *Journal of Political Economy*, Vol. 99(5), pp. 899-928

[16] Bernheim, B. D. and Levin, D. K. 1989, Social Security and Personal Saving: An Analysis of Expectations, *The American Economic Review*, Vol. 79, No. 2, pp. 97- 102

[17] Bernheim, B.D., Shleifer, A. and Summers, L., 1985, Strategic Bequest Motive, *The Journal of Political Economy*, Vol. 93(6), pp. 1045-1076

[18] Bertrand, M. and Mullainathan, S. 2001, Do People Mean What They Say? Implications for Subjective Survey Data,*The American Economic Review*, Vol. 91, No. 2, pp. 67-72

[19] Bisin, A. and T. Verdier, 2001, The Economics of Cultural Transmission and the Dynamics of Preferences, *Journal of Economic Theory*, 97(2), pp. 298-319

[20] Brandt Lorent, Jikun Huang, Guo Li and Scott Rozelle, 2002, Land Rights in Rural China: Facts, Fictions and Issues, *The China Journal*, Vol. 47, pp. 67-97

[21] Brock, W. and Durlauf, S., 2001, Interactions-based Models in *Handbook of Econometrics* edited by Heckman and Leamer, vol. 5, pp.3297-3380, Amsterdam: North Holland

[22] Brown, J. R., and Finkelstein, A., 2008, The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market, *American Economic Review*, 98(3): pp. 1083-1102

[23] Bowlus A.J., T. Sicular, 2003, Moving toward markets? Labor allocation in rural China, *Journal of Development Economics*, 71, pp. 561583

[24] Chan, K.W., 2010, The Household Registration System and Migrant Labor in China: Notes on a Debate, *Population and Development Review*, Vol. 36(2), PP. 357364

[25] Chan, K.W., and Zhang, L., 1999, The Hukou System and Rural-Urban Migration in China, *China Quarterly*, No. 166, pp. 818-855

[26] Chan, S. and Stevens, A., 2004, Do changes in pension incentives affect retirement? A longitudinal study of subjective retirement expectations, *Journal of Public Economics*, Vol. 88, Iss. 7-8, pp. 1307-1333

[27] Chen, Shaohua, Ren Mu, and Martin Ravallion, 2009, Are there lasting impacts of aid to poor areas?, *Journal of public economics*, 93(3), pp. 512-528.

[28] Cohen, A., and Einav,L., 2007, Estimating Risk Preferences from Deductible Choice, *American Economic Review*, Vol. 97(3), 745-88

[29] Cutler, D., Finkelstein, A. and McGarry, K., 2008, Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance, *American Economic Review*, Vol. 98(2), 157-162

[30] de Finetti, B. 1969, The True Subjective Probability Problem Versus Various More or Less Related Side Problems, *Proceedings of a Research Conference on Subjective Probability and Related Fields*, Psychological Institute of the University of Hamburg, pp. 33-42

[31] Dixit, A., 2003, Trade Expansion and Contract Enforcement, *Journal of Political Economy*, Vol. 111, pp. 1293-1317

[32] Dominitz, J. and Manski, C. F. 1994, Using Expectations Data to Study Subjective Income Expectations, *Econometrica*, pp. 941-1003

[33] Dominitz, J., Manski C. F. and Heinz, J. 2002, Social Security Expectations and Retirement Savings Decisions, NBER working paper 8718

[34] Durlauf,S., 2001, Framework for the Study of Individual Behavior and Social Interactions, *Sociological Methodology*, Vol. 31, pp. 47-87

[35] Easterly, William and Levine, Ross, 1997, Africaś Growth Tragedy: Policies and Ethnic Divisions, *Quarterly Journal of Economics*, 112(4), pp. 1203-1250

[36] Einav, L. *et al.* , How general are risk preferences? Choices under uncertainty in different domains, *American Economic Review*, Forthcoming

[37] Engle-Warnick, J. and R. L. Slonim, 2003, Learning to Trust in Indefinitely Repeated Games, working paper

[38] Esteban, Joan and Ray, Debraj, 1994, On the Measurement of Polarization, *Econometrica*, 62(4), pp. 819-51

[39] Fang, H., Keane, M. and Silverman, D. 2008, Sources of Advantageous Selection: Evidence from the Medigap Insurance Market, *Journal of Political Economy*, Vol.116, No. 2, pp. 303-49

[40] Finkelstein, A. and MaGarry, K., 2006, Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market, *American Economic Review*, Vol. 96(4), pp. 938-958

[41] Francois, P., 2002, Social Capital and Economic Development, Newyork, Routledge

[42] Francois, P. and van Ypersele, T., 2009, Doux Commerces: Does Market Competition Cause Trust?, working paper, The Univ. of British Columbia

[43] Francois. P. and Zabojnik, J., 2005, Trust, Social Capital and Economic Development, *Journal of the European Economics Association*, 3(1), 51-94

[44] Gigerenzer, G. 1991, How to make cognitive illusions Disappear: Beyond Heuristics and Biases, in *European Review of Social Psychology* editted by Stroebe and Hewstone, Vol. 2, John Wiley& Sons Ltd.

[45] Gleaser, E.,Laibson, D., Scheinkman, J., and Soutter,.D, 2002, Measuring Trust, *Quarterly Journal of Economics*, Aug. 2000, pp. 811-846

[46] Goda, G. S., 2010, The Impact of State Tax Subsidies for Private Long-Term Care Insurance on Coverage and Medicaid Expenditures, Working Paper

[47] Gordon, R., Franklin, N. and Beck, J. 2005, Wishful thinking and source monitoring, *Memory & Cognition*, Vol. 33, pp. 418-429

[48] Government Accountability Office (GAO), 2005, Long-term care financing, Washington, DC

[49] Grief, A., 1994, Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies, *Journal of Political Economy*, 102, 912-950

[50] Grove,W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., and Nelson, C. 2000, Clinical versus mechanical prediction: A metaanalysis, *Psychological Assessment*, Vol. 12, pp. 19-30

[51] Guiso, L., Sapienza, P. and Zingales, L., 2004, The Role of Social Capital In Financial Development, *The American Economic Review*, Vol 94(3), pp. 526-556

[52] Guiso, L., Sapienza, P. and Zingales, L., 2009, Cultural Biases in Economic Exchange, *The Quarterly Journal of Economics*, pp. 1095-1131

[53] Hall, A.R., Rudebusch, G.D., and Wilcox, D.W., 1996, Judging Instrument Relevance in Instrumental Variable Estimation, *International Economic Review*, 37, 283-289

[54] Ham John, John Kagen and Steven Lehrer, 2005, Randomization, Endogeneity and Laboratory Experiments: The Role of Cash Balances in Private Value Auctions, *Journal of Econometrics*, Vol 125(1-2), pp. 175-205

[55] Hamermesh, D. 1985, Expectations, Life Expectancy, and Economic Behaviour, *The Quarterly Journal of Economics*, Vol. 100, No. 2, pp. 389-408

[56] Hausman, J. A. 1978, Specification tests in econometrics, *Econometrica*, Vol. 46, pp. 1251-1271

[57] Health Insurance Association of America, 2000, Long term care Insurance in 1997 -1998, Washington D.C.

[58] Heckman, J.J., 1981, The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process, in *Structural Analysis of Discrete Panel Data with Econometric Applications*, ed. by C. Manski and D. McFadden, Cambridge: MIT Press

[59] Hertzog, C. 1989. Influences of cognitive slowing on age differences in intelligence, *Developmental Psychology*, Vol. 25, pp. 636-651.

[60] Hoffrage, U., Lindsey, S., Hertwig, R. and Gigerenzer, G. 2000, Communicating Statistical Information, *Science*, Vol. 290, Iss. 5500, pp. 2261-2262

[61] Hogarth, R. 1975, Cognitive Processes and the Assessment of Subjective Probability Distributions, *Journal of the American Statistical Association*, Vol. 70, No. 350 , pp. 271- 289

[62] Holden, K., McBride, T. and Perozek, M. 1997, Expectations of Nursing Home Use in the Health and Retirement Study: The Role of Gender, Health, and Family Characteristics, *Journal of Gerontology: social science*, Vol. 52B, No. 5, pp. 240-251

[63] Huck, Lunser and Tyran, 2007, Competition fosters trust, working paper, University College of London

[64] Hurd, M., 1987, Savings of the Elderly and Desired Bequests, *The American Economic Review*, Vol. 77(33), 298-312

[65] Hurd, M. and McGarry, K., 1995, Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study, *The Journal of Human Resources*, Vol. 30, Special Issue on the Health and Retirement

[66] Hurd, M. and McGarry, K. 2002, The Predictive Validity of Subjective Probabilities of Survival, *The Economic Journal*, Vol. 112, No. 482, pp. 966-985

[67] Jing, J., 1996, The Temple of Memories: History, Power, and Morality in a Chinese Village, Standford University Press

[68] Johnson, Elizabeth L., 1976, Households and Lineages in a Chinese Urban Village, Ph.D. dissertation, Cornell University

[69] Jurges, H., 2001, Do Germans Save to Leave an Estate? An Examination of the Bequest Motive, *Journal of Scand. Economics*, 103(3), 391-414

[70] Juster, F. T., 1966, Consumer Buying Intension and Purchase Probability: An Experiment In Survey Design, *Journal of the American Statistics Association*, Vol. 61(315), pp. 658-96

[71] Kahneman, D. and Frederick, S. 2002, Representativeness revisited: Attribute substitution in intuitive judgment, in *Heuristics of Intuitive Judgment: Extensions and Applications* edited by T. Gilovich, D. Griffin and D. Kahneman, New York: Cambridge University Press

[72] Kemper, P., Spillman, B.C., Murtaugh, C.M., 1991, A lifetime perspective on proposals for financing nursing home care, *Inquiry*, Vol. 28, pp. 333-344

[73] Knack, S., and Keefer,P., 1997, Does Social Capital Have an Economic Payoff? A Cross Country Investigation, *Quarterly Journal of Economics*, pp. 1251-1288

[74] Koneztka, R. T. and Luo, Y., 2010, Explaining Lapse in Long Term Care Insurance Markets, *Health Economics*, forthcoming

[75] Kopczuk, W. and Lupton, J., 2007, To Leave or Not to Leave: The Distribution of Bequest Motives, *The Review of Economic Studies*, pp. 207-236

[76] LaPorta, R., Lopez-de-Salanes, F., Shleifer, A., and Vishny, R., 1997, Trust in large organizations, *American Economic Review Papers and Proceedings*, Vol. 87, pp. 333–338

[77] Manapata, M., Nowaka, M., and Randa, D. [2012], Information, irrationality, and the evolution of trust, *Journal of Economic Behavior and Organization*, Forthcoming

[78] Manski, C., 1993, Identification of Endogenous Social Effects: the Reflection Problem, *Review of Economic Studies*, Vol. 60, pp. 531-542

[79] Manski, C., 2004, Measuring Expectation, *Econometrica*, pp. 1329-1376

[80] Meehl, P. E. 1954, Clinical versus statistical prediction, Minneapolis: University of Minnesota Press

[81] Miguel, E. and Gugerty, M., K., 2004, Ethnic Diversity, Social Sanctions and Public Goods in Kenya, *Journal of Public Economics*, 89, pp. 2325-2368

[82] Miguel, E., Gertler, P. and Levine, D., 2005, Does Social Capital Promote Industrilization? Evidence From a Rapid Industrializer, *The Review of Economics and Statistics*, 87(4), pp. 754-762

[83] Morin, R. A. and Suarez, A. F. 1983, Risk aversion revisited, *The Journal of Finance*, Vol. 38, pp. 1201-1216

[84] Narayan, D. and Prichett, L., 1999, Cents and Sociability: household income and social capital in rural Tanzania, *Economic Development and Social Change*, Vol. 47(4), pp. 871-897

[85] Naughton, B., 2007, The Chinese Economy: Transitions and Growth, Cambridge: MIT Press

[86] Nigel, H. 1992, Wishful thinking impairs belief-desire reasoning: A case of decoupling failure in adults, *Cognition*, Vol. 45, pp. 141-162

[87] Nyarko, Y. and Schotter, A., 2002, An Experimental Study of Belief Learning Using Elicited Beliefs, *Econometrica*, Vol. 70(3), pp. 971-1005

[88] Papka, L. and Wooldridge, J. 1996, Econometric Methods for Fractional Response Variables With an Application to 401 K Plan, *Journal of Applied Econometrics*, Vol. 11, No. 6, pp. 619-632

[89] Pauly, M., 1990, The Rational Nonpurchase of Long Term Care Insurance, *Journal of Political Economy*, Vol. 98, No. 1, 153-168

[90] Platteau, J., 2000, Institutions, Social Norms, and Economic Development, Harwood Academic Publishers, Amsterdam

[91] Putnam R., R. Leonardi, and R. Nanetti, 1993, Making Democracy Working: Civic Tradition and Modern Italy, Princeton: Princeton University Press

[92] Ruthanna, G., Franklin, D. and Beck, J. 2005, Wishful thinking and source monitoring, *Memory & Cognition*, Vol. 33, pp. 418-429

[93] Salthouse, T.A., and Mitchell, D. R. D. 1990, Effects of age and naturally occurring experience on spatial visualization performance, *Developmental Psychology*, Vol. 26, pp. 845-54

[94] Savage L. 1954, Foundation of Statistics, Dover Publications

[95] Schechter, Laura, 2007, Theft, Gift-Giving, and Trustworthiness: Honesty Is Its Own Reward in Rural Paraguay, *American Economic Review*, pp.1560-1582

[96] Shaw, K. 1996, An Empirical Analysis of Risk Aversion and Income Growth , *Journal of Labour Economics*, Vol. 14, No. 4, pp. 626-653

[97] Shein, Louisa, 1994, The dynamics of cultural revival among the Miao in Guizhou, *New Asia Academic Bulletin*, pp. 199-212

[98] Smith, K., Taylor, D. and Sloan, F. 2001, Longevity Expectations and Death: Can People Predict Their Own Demise, *The American Economic Review*, Vol. 91, No. 4, pp. 1126-1134

[99] Soceity of Acturies, 2007, Long term care experience committee intercompany study 1984-2004, Report, Schaumburg, Illinois

[100] Solow, R., 1999, Notes on Social Capital and Economics Performance in *Social Capital: A Multifaceted Perspective* edited by Dasgupta and Serageldin, World Bank, pp. 6-9

[101] Strauch, J., 1983, Community and Kinship in Southeastern China: The View from the Multilineage Villages of Hong Kong, *Journal of Asian Studies*, 44(1), pp. 21-50

[102] Stokey, N., Lucas, R.E., with Prescott, E., 1991, Recursive Methods in Economic Dynamics, Harvard University Press, Cambridge, M.A.

[103] Stewart *et al.*, 2009, Annual expenses for Nursing Home Care, *Medical Care*, Vol. 47, No. 3, pp. 295-301

[104] Sung, J. and Hanna, S. 1996, Factors related to risk tolerance, *Financial Counseling and Planning*, Vol. 7, pp. 11-20

[105] Tabellini, G., 2008, The Scope of Cooperation: Values and Incentives, *The Quarterly Journal of Economics*, pp. 905-950

[106] Taylor D. H., Ostermann, J., Will Acuff, S. and Ostbye, T. 2005, Do Seniors Know Their Risk of Moving to a Nursing Home, *Health Services Research*, Vol. 40, pp. 811-828

[107] Tomes, N., 1981, The Family, Inheritance, and the Intergenerational Transmission of Inequality, *The Journal of Political Economy*, Vol. 89, pp. 928-58

[108] Tourangeau, R. 1984, Cognitive sciences and survey methods, in *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines* edited by Labine, Loftus, Straf, Tanur and Tourangeau, National Academy Press, Washington, DC

[109] Tversky, A. 1974, Assessing Uncertainty, *Journal of the Royal Statistical Society*, Ser. B, 36, pp. 148-159

[110] Tversky, A., and Kahneman, D. 1971, The Belief in the Law of Small Numbers, *Psychological Bulletin*, Vol. 76, pp. 105-110

[111] Tversky, A., and Kahneman, D. 1974, Availability: A Heuristic for Judging Frequency and Probability, *Cognitive Psychology*, Vol. 5, pp. 207-232

[112] Tversky, A., and Kahneman, D. 1981, The Framing of Decisions and the Psychology of Choice, *Science*, pp. 453-458

[113] Tversky, A., and Kahneman, D. 1985, Judgment Under Uncertainty: Heuristics and Biases, *Science*, pp. 1124-1131

[114] Vigdor, J. L., 2002, Interpreting Ethnical Fragmentation Effects, *Economics Letters*, 75, pp. 271-276

[115] Wallsten, T. and Budescu, D. 1983, Encoding Subjective Probabilities: A Psychological and Psychometric Review, *Management Science*, Vol. 29, No. 2, pp. 151-173

[116] Wang H. and Hanna S. 1997, Does Risk Tolerance Decrease With Age, *Financial Counseling and Planning*, Vol. 8, pp. 27-31

[117] Weesie, J. 1999, Seemingly unrelated estimation and the cluster-adjusted sandwich estimator, *Stata Technical Bulletin*, Vol. 32, pp. 34-47

[118] White, H. 1982, Maximum likelihood estimation of misspecified models, *Econometrica*, pp. 1-25

[119] White, H. 1996, Estimation, Inference and Specification Analysis. Cambridge: Cambridge University Press

[120] Woon, Y., Social Change and Continuity in South China: Overseas Chinese and the Guan Lineage of Kaiping County, 194987, *The China Quarterly*, 118, pp. 324-344

[121] Zhao, Yaohui, 1999, Leaving the countryside: rural-to-urban migration decisions in China, *The American Economic Review*, vol. 89(2), pp. 281-286

[122] Zhang, Kevin Honglin, and Shunfeng Song, 2003, Ruralurban migration and urbanization in China: Evidence from time-series and cross-section analyses, *China Economic Review*, Vol. 14(4), pp. 386-400