INFORMATION TO USERS

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

RISKS, INSURANCE, AND THE LOSS OF WORK DUE TO ILLNESS AMONG LOW-INCOME HOUSEHOLDS

by

RONG YI

B.S., Peking University, 1995
M.A., Boston University, 2000

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2002

UMI Number: 3015609

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Approved by

First Reader ______________________
Randall P. Ellis, Ph.D.
Professor of Economics

Second Reader ______________________
Maristella Botticini, Ph.D.
Assistant Professor of Economics

Third Reader ______________________
Robert E. B. Lucas, Ph.D.
Professor of Economics

ACKNOWLEDGEMENTS

Research is like a navigation, throughout which guide-lights from keystone towers help greatly in a successful journey. I started working on this dissertation since the fall of 1999. It has been a continuous exploration made by me and shared by many others. My special gratitude goes to Randall Ellis, my advisor, mentor, and good friend, who has offered me invaluable helps and insights during my dissertation research and throughout my graduate studies. I would also like to thank Maristella Botticini and Dilip Mookherjee, who have offered great inputs into my research. Many thanks to my family, whose love and support have given me the courage and determination to continue on at the various times of difficulties. Many thanks to Robert Lucas, Marc Rysman, Peter Doerginer, my friends at the Department of Economics at Boston University, and other seminar participants for their valuable comments and suggestions. I assume the full responsibility of all errors and mistakes.

RISKS, INSURANCE, AND THE LOSS OF WORK DUE TO ILLNESS AMONG LOW-INCOME HOUSEHOLDS

(Order No.   )

RONG YI

Boston University Graduate School of Arts and Sciences, 2002

Major Professor: Randall P. Ellis, Professor of Economics

ABSTRACT

For low-income households in developing countries, formal financial markets usually do not function well due to asymmetric information and lack of enforcement. Informal risk-coping strategies often arise, compensating for market failures. Using a large household survey data set from southern India, the ICRISAT data, this dissertation examines the effectiveness of village-level mutual insurance, and the interactions between consumption, health, and labor supply.

The dissertation starts with a review of the literature on informal risk-coping strategies, and offers some insights on the methodologies and empirical findings in this area. The socioeconomic conditions of the ICRISAT households are discussed afterwards.

The theoretical analyses compare the predictions on consumption and sickness leave under full insurance for three preference specifications—additive CARA, multiplicative CRRA, and nonseparable preferences. It shows that the first two specifications offer very

similar predictions, i.e., individual consumption and sickness leave can be fully explained by the village aggregates and a set of individual factors. Sickness leave does not affect consumption and vice versa. Under nonseparability between consumption and health, however, individual consumption and sickness leave interact after controlling for village aggregates and individual factors, which is still compatible with full insurance.

The empirical analyses of this dissertation consist of three chapters. The first chapter addresses the problem of missing observations in the labor data. The results suggest serious selection biases towards low-income households. A selection factor is constructed and brought to empirical estimations later. The second chapter revisits the famous work of Townsend and tests full insurance with more years of data, village-specific prices, and sample selection adjustment. The results strongly reject full insurance and also suggest serious biases in Townsend's regressions. The last chapter tests full insurance and nonseparability structurally. The results reject full insurance and separability between consumption and health on the grounds that socioeconomic factors such as household income, landholding, and demographic characteristics which may restrict the scope and coverage of mutual insurance affect individual consumption and sickness leave. The results also suggest that the ICRISAT households under-invest in health by taking less sickness leave and more consumption, other things equal. Policy implications are discussed afterwards.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CARA: Constant Absolute Risk Averse

CRRA: Constant Relative Risk Averse

CV: Covariance of Variation

ICRISAT: International Crop Research Institute for Semi-Arid Tropics

PIH: Permanent Income Hypothesis

CHAPTER 1 LITERATURE OVERVIEW

One of the salient features of the market system in developing countries is the lack of information and enforcement mechanisms, which in turn cause failures in important financial institutions such as credit and insurance. This chapter provides an overview of major research themes in this area and focuses on Townsend's work (1994) since it is closely related to my research.

1.1 RISKS AND RISK-COPING STRATEGIES IN DEVELOPING COUNTRIES

Credit and insurance are the two most important risk-coping strategies individuals and households resort to under uncertainties. Credits allow people to accumulate and deplete financial or nonfinancial assets across time, whereas insurance allows people to allocate economic resources and activities across the states of nature. Pareto optimal allocation of economic resources can be achieved under certain premises when there is full credit and insurance. Assuming certain separability in preference, perfect credit implies that the marginal utility of consumption in each time period is based solely on the expected lifetime income, while perfect insurance implies that the marginal utility of consumption is the same in all possible states regardless of the realization of the uncertainties. The former is usually referred to as the permanent income hypothesis and the latter consumption smoothing.

Household income is usually low and volatile in the rural areas of developing countries largely due to the important role played by agriculture. Rural households are vulnerable to factors such as weather fluctuations, commodity price changes, business and crop failures, and human illnesses. For example, the estimated coefficient of variation in income is around 40 percent in the South Indian villages surveyed by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). When considering farm profits only, the coefficient is found to be over 125 percent.

How do such large income fluctuations affect welfare? One approach is to gauge the willingness to pay to completely eliminate income variability under a certain preference specification. The following illustrates the use of this framework with constant relative risk-averse preferences.

A representative household receives an uncertain income stream y with mean Y and variance σ^2. Let m denote the amount of money the household is willing to pay so that the utility from receiving $Y-m$ with certainty is equal to the expected utility from receiving the uncertain income. $Y-m$ is sometimes referred to as the "certainty equivalence" of income stream y. m solves

$$(1.1) \quad U(Y-m)=E[U(y)].$$

A second-order Taylor expansion around Y yields

$$(1.2) \quad U(Y)-mU'(Y)+\frac{m^2}{2}U''(Y)\approx E[U(Y)]+E[U'(Y)(y-Y)]+E\left[\frac{m^2}{2}U''(Y)\right].$$

Assuming constant relative risk-averse of degree *R*, where

$$(1.3) \quad R=\frac{-YU''(Y)}{U'(Y)},$$

we have

$$(1.4) \quad U''m^2-2U'm-U''\sigma^2\approx 0.$$

Discarding the negative root, we have

$$(1.5) \quad \frac{m}{Y}\approx -\frac{1}{R}+\sqrt{\frac{1}{R^2}+\sigma^2}.$$

For a moderate relative risk aversion $R = 2$ and a coefficient of variation of 0.4, households would be willing to give up approximately 14 percent of their income to ensure perfect income stability. This represents a large absolute amount for a farm household living close to subsistence levels. Of course, these calculations will overstate the importance of risks and the willingness to pay for protections if there already exists some form of smoothing.

As these calculations illustrate, the need for protection from risks is significant. However, in developing countries, especially in the rural areas, formal financial markets usually do not exist, and for those that do, many work imperfectly. The primary reasons are the lack of information and limits to contractual enforcement. Let us consider a credit market in a rural area in a developing country. A lender wants to at least cover his costs of lending, even if not for profit. The problem is that some borrowers are intrinsically bad risks, but the lender is not able to distinguish them from the good risks either due to the lack of information on their credit history and financial background, or because the costs of acquiring such information are high. Both moral hazard and adverse selection will be severe. The borrowers will always have the incentive to overinvest in riskier projects. In addition, because of limited liability, the lender will be afraid of increasing the interest rate at the cost of attracting more riskier borrowers, or increase the volume of loans because lending more may lead to higher losses to failed projects. The amount of loans and the coverage of protection against risks are limited, resulting in welfare losses.

The problems with insurance are similar, however, there is a particular problem in developing countries that restricts the insurability of risks. Full insurance is based on the presumption that risks are idiosyncratic and can be fully pooled. In reality, however, risks tend to be correlated at both individual and community level. Random factors such as rainfall, humidity, temperature, crop diseases, pests, and commodity prices are highly covariate locally. A group of households from the same village will be unable to provide mutual insurance for each other against covariate risks if they rely on the same mechanisms for risk protections at the same time. In circumstances like this, village level buffer stocks are usually used for smoothing purposes. For example, Lim and Townsend

(1994), and Paxson and Chaudhuri (1994) show that the ICRISAT households' primary device in smoothing village-level aggregate shocks is crop inventory.

The more pertinent question regarding insurance, perhaps, is whether there exists a meaningful scope of mutual insurance given the local specifics. In particular, one needs to decompose the sources of income and consumption fluctuations and determine whether or not the dominant source of fluctuations is highly correlated across individuals and households. Townsend (1994) finds that household incomes do not comove across at the village level even though all households have to cope with the highly covariate rainfall uncertainties. He considers this as showing the potential for mutual insurance and argues that it is because "...households earn their income in different ways, subject to different risks, and do not diversify much".

Economic theories predict that households with access to improved financial system can use credit and insurance to diversify risks and potentially to raise income, consumption and savings (see Ray, 1998). Conversely, Greenwood and Jovanovic (1990), and Banerjee and Newman (1993) develop models in which entry costs and poorly functioning financial systems can limit access by the poor, reinforcing low levels of income, consumption and savings for certain groups. Thus a lack of credit and insurance can contribute to poverty and inequality in society as a whole. Moreover, imperfect markets are contagious. As many have pointed out, market failures in some sector can lead to problems in other markets as well (see Ray, 1998).

Given the size of the risks and the magnitude of welfare losses, it will not be surprising to observe various informal risk-coping mechanisms arising filling in the holes left by market failures. The information flow, contractual relationship, and the scope and efficiency of such informal risk-coping strategies have been intensively studied in the past fifteen years or so in the literature of development economics. I focus here on three major data sets that have been closely examined in the empirical research on informal risk-coping strategies. They are, the three South Indian villages sampled by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), the set of counties constituting regional economies and national economy of Thailand sampled by the Thai Socio-Economic Survey (SES), and the set of villages constituting the rural economy of the Cote d'Ivoire sampled by the World Bank's Living Standards Measurement Section (LSMS). Numerous studies have used the three data sets, and special estimation strategies been employed to incorporate the unique features of the data. Townsend (1995) has offered some general discussions on the merits and drawbacks of each of the three data sets.

The ICRISAT data is a panel data set, which traces 120 households from three villages in the semi-arid tropical India continuously for ten crop years. Information on household demographic composition, labor supply, input and output by plot, household financial accounts and transactions, monthly commodity prices and rainfall, are included among other details. It contains tractable accounts of major household economic activities over a continuous times series. The mostly notable drawback of this data is its consumption measures, which is not directly measured but estimated as a residual. There are selection issues as well.

The Thai SES data is highly regarded as providing rather reliable estimates of consumption and income, including home-produced goods. There are numerous quality control checks, such as repeat interviews triggered by an inability to balance the monthly accounts of income and expenditures to within 10 percent. The major drawback of the Thai data is that it is a repeated survey for 5 nonconsecutive years spanning 1975 to 1990, despite the fact that the sampled households might have been with the survey all the time.

The Cote d'Ivoire LSMS data consists of two overlapping household panels in 1985-86 and 1986-87. Clusters or villages were the primary sampling unit, and these were spread throughout the three major geographic divisions of the country. Deaton (1994) provides an excellent discussion of the reliability of the LSMS consumption and income modules. One of the most impressive achievements of the LSMS data is its effectiveness in data collection with the help of computer technology. Responses were quickly taken to local headquarters, and entered into PCs, and then immediately run through editing programs, so that cross-checks and corrections could be carried out on subsequent visits.

There are two major themes in empirical studies on these survey data. One theme focuses on the role of informal insurance in smoothing income before shocks occur, such as sharecropping, crop-plot portfolio management, and labor diversification (see

Morduch, 1995, and Townsend, 1995). The other theme focuses on the role of informal credits in smoothing consumption after shocks are realized, such as accumulation and depletion of productive assets, sales of stored produce, and transfer incomes in social networks (see Lucas and Stark, 1988, Eswaran and Kotwal, 1989). Theoretically speaking, consumption smoothing and income smoothing are perfectly substitutable for the purpose of risk protection if markets are perfect. In reality, however, households are found to adopt many different approaches according to the specific local conditions.

Some studies focus on the nature of the smoothing strategies and their effectiveness. While looking at the local communities, it is justifiable to assume certain smoothing channels do not exist, so that we can theoretically examine a particular form of risk protection without loss of generality. However, empirical inferences based on the theoretical framework might have over-estimated the effectiveness of a particular form of smoothing, because other smoothing channels contributing to risk protections are assumed away. Policy implications therefore should also be taken with great cautions.

Starting with different assumptions about the local conditions, empirical studies have reached quite different results from the same data. For example, in a research on income smoothing, Rosenzweig and Wolpin (1993) finds that the ICRISAT villagers have few financial assets and tend to resort to mobile nonfinancial assets such as bullocks in coping with income shortfalls. In particular, they find that over 86 percent of the sample households reported purchases or sales of bullocks, and the sales of bullocks were 3.5

times as prevalent as the sales of jewelry. They then formulate a dynamic setting in which agricultural investment is determined under income uncertainty and borrowing constraints. Their estimations indicate that net purchases are significantly more likely to occur when income is high than when income is low. Using the same data, Kochar (1995, 1999) argues that the ICRISAT households are able to protect themselves from income variations by compensating increases in market hours of work, which reduces the need to resort the depletion of productive assets. In particular, she finds that households shift from own-farm cultivation to the labor market when crop profit drops. She also points out that the ability of smoothing income through labor diversification is mitigated by demographic shocks such as sickness, death, and the dissolution of household, and under circumstances like these, households increase informal borrowing.

There is evidence that households also engage in social arrangement such as migration, marriage and remittances for income smoothing purposes. In the study of informal risk-sharing arrangements among family members across the ICRISAT villages, Rosenzweig (1988) finds that households seem to marry daughters deliberately out over space, in order that remittance can flow between areas depending on which side is suffering a negative income shock. In the Cote d'Ivoire, Grimand (1992) sorted the LSMS data by tribe and shows that networks among tribes allow greater consumption smoothing than is apparent in the nonsorted data. Paulson (1994) finds remittances from migrants help stabilize household consumption in the highly migrating Thai SES communities.

In addition to *ex ante* income smoothing, households engage in *ex post* consumption smoothing. For example, Lim and Townsend (1994), and Paxson and Chaudhuri (1994) find that crop inventory and currency are used for consumption smoothing among the ICRISAT households. Storage of crops has been used since biblical time, although holding crop stocks can be costly because they are not exactly durable. These consumption smoothing techniques are preferred to savings and credits if rural banks are few or far, or if there are restrictions on the rapid liquidation of savings. Another important *ex post* risk-coping channel is credit. Based on a yearlong survey in northern Nigeria, Udry (1994) shows that informal credit contracts play a direct role in pooling risks between households. He develops two models of state-contingent loans where borrowers' repayments depend on the realization of random shocks by both borrowers and lenders. The first is a competitive equilibrium with perfectly enforceable contracts. The second permits imperfect information and equilibrium default. Estimates from both models indicate that quantitatively important state-contingent payments are embedded in the loans, but a fully efficient risk-pooling equilibrium is not achieved.

I would like to point out that without specific information on the actual strategies households adopt, we could at most say that these observed informal smoothing strategies all have played a role. No villages are alike. In fact, the actual smoothing approaches households have adopted are much more diverse and complex than assumed in theoretical works. For example, Paulson (1994) studies the role of migration and remittances in income and consumption smoothing in Thailand using the Thai SES data. She was able to conduct a field research in the study villages focusing on the differences between good years and bad years and on whether or how households might have managed to smooth consumption (also see Townsend, 1995). She finds a flexible range of smoothing patterns. For instance, some relatively rich households try to smooth consumption with buffer stock, while the relatively poor households achieve the same goal by increasing labor supply. In addition, there exists organizations and networks such as rice banks, housewife funds, or even health insurance fund for those who are more connected to the local society. These institutions appear to have contigencies in loan contracts and direct transfers to those in distress. She also finds that market conditions vary markedly across villages even within the same vicinity. For example, one village in the Maajaam county is replete with village institutions and is very organized. Another village in the same county appears to be less successful in credit and insurance arrangements. The third has virtually no institutions. These observations suggest a very important thing, that is, researchers should be leery of generalizations about what villages can or cannot do.

While evaluating the effectiveness of the informal smoothing strategies, empirical tests are built on the theory of complete markets. The community under study is usually modeled as a closed economy with internal credit or insurance of a certain kind. If there are perfect credit markets, consumption of a representative individual or household is not affected by transitory shocks over time and only comoves with his expected lifetime earnings. Thus, if we run a regression with individual or household consumption as the dependent variable and transitory income as one of the explanatory variables, the more complete are the credit markets, the closer the coefficient on transitory income will be to zero. Similarly, according to the theory of full insurance, individual or household consumption should follow community average and nothing else, and movements in individual or household consumption is fully explained by aggregate shocks not smoothed at individual level. Therefore, in the regression of individual or household consumption on community average consumption, household income and other factors, the more complete are insurance, the closer the coefficient on community average consumption will be to 1, and the coefficient on household income to zero.

The ICRISAT and Cote d'Ivoire data contain household-level information as well as exogenous community-level information such as rainfall. Empirical tests can be constructed almost directly from the theories. For the Thai data, information on particular households is not available, and the common approach is to compare various subgroups with a larger aggregate instead.

Previous empirical findings suggest that households are able to smooth income and consumption considerably, but complete protection against risks is not achieved. In addition, some households are better protected than others. For example, in a test for complete credit markets, Paxson (1992) estimates with the Thai data that the coefficient on transitory income is between 0.73 and 0.83, but not significantly different from 1. Townsend (1994) tests for full insurance with the ICRISAT data and reaches quite similar estimates. He also finds evidence that the landless are less well insured.

There are good reasons why observed economies achieve something less than full insurance. As discussed earlier, informal credit and insurance systems are eroded by counterproductive incentives such as moral hazard and adverse selection. To better understand the actual environment, a theoretical portrait should be able to incorporate more micro-level contractual structures such as principal-agent relationships and enforcement mechanisms in general equilibrium. Calculations involved in such models are mammoth, nonetheless there has been great promising steps in this area (Phelan and Townsend, 1991, Green, 1987, Spear and Strivastava, 1987).

1.2 HEALTH AND HEALTH RISKS IN DEVELOPING COUNTRIES

1.2.1. HEALTH AS A CONSUMPTIVE GOOD AND A CAPITAL GOOD

Health is one of the most important welfare indicators, and the role it plays is multi-folded. Good health is a highly valued asset itself, and perfect health is sometimes viewed as a target. It is the prerequisite for many activities such as enjoying consumption and leisure. It also affects one's ability of earning incomes from the labor market. There has been empirical evidence on the positive relationship between health/nutrition and wage rate, particularly in developing countries (see Ray, 1998, and Deolalikar, 1988).

Just a little more than two centuries ago, health was generally viewed as fate. Good health was considered a gift from God and bad health was merely bad luck. Nowadays it is portrayed as a capital good due to the success of modern medicine. Individuals can affect their health status effectively and are able to achieve a reasonable health state through proper means. In theoretical health economics, the idea of health as a capital good has been well developed in dynamic settings. *Health capital* is usually formulated as depreciating over time, receiving random nonpositive shocks such as illness or injury, and renewed by "investment" in the form of medical care and other efforts aiming to improve health (See Grossman 1972, 2000, Zweifel and Breyer 1997).

In addition to good health, there are other objectives in life. People usually find themselves facing the tradeoff between health and the other aims, for instance, additional income from overtime work. One might wonder what is the best allocation between health and other goals in life from an individual's point of view. Following the conventions of consumer theories, we may translate this into preferences over health and other goods, in particular, consumption. The optimization problem can be stated as follows:

An individual derives utility from consumption C and health H, which in turn are obtained from consumer goods X by $C = C(X)$ and medical care M by $H = H(M)$. Let p and q denote the market price for consumer goods and medical care, and we can write the budget constraint as $Y = pX + qM$. Individual optimization is achieved by choosing an allocation (X^*, M^*) of consumer goods and medical care so that utility is maximized with the given budget constraint, and the corresponding consumption and health status is (C^*, H^*).

If health status is fixed, as assumed by many works implicitly or explicitly, medical care does not have any impact on health. $H(M)$ degenerates to a fixed level, H^o. Income Y becomes exogenous and the budget constraint is also independent of medical care. The optimum will be $M = 0$ and $X = Y/p$, with the corresponding consumption level at C^o, and indifference curve U^o.

When health status is variable and positively related with medical care, $H(M)$ becomes concave, i.e., health improves with medical care spending, but additional gains in health reduce as medical care spending increases. When income also depends on health, the substitution between consumer goods and medical care is no longer linear, and the budget constraint becomes $Y(M) = pX + qM$. Reductions in medical spending do not lead to reductions in health or income as much, and the gains to increased consumption are more than enough to compensate the utility losses associated with health deterioration along the indifference curve. However, the rate of transformation may not be uniform. There may exist a turning point, where further reduction in medical spending will lead to larger and larger drops in health and income, so that the budget constraint bends toward the origin. Additional gains to higher consumption are no longer enough to compensate the utility losses in health along the indifference curve. Therefore, the turning point is the optimal allocation, with corresponding consumption and health at (C^*, H^*).

An important part of this dissertation examines the demand for consumption and health under health risks and village-level mutual insurance. The above discussions are very helpful and intuitive for us to understand the relationship between consumption and health under perfect insurance and imperfect insurance. If the village-level mutual insurance is perfect, health risks are i.i.d., and there is no aggregate risks at the village-level, the budget constraint a representative individual faces degenerates to a fixed level because of risk-sharing at the village level. If insurance is imperfect, the individual faces a curved budget constraint, and chooses (C^*, H^*) at the turning point to maximize utility.

Controlling for the severity of health problems, it is easy to show that individuals with higher income will be on higher indifference curves than otherwise.

1.2.2 HEALTH AND HEALTHCARE IN DEVELOPING COUNTRIES: MAJOR THEMES AND MAJOR FINDINGS IN EMPIRICAL RESEARCH

Health is particularly important to developing countries not only because of its consumptive and income benefits, but also because of its dynamic impacts on income, consumption and welfare. Unlike the usual capital good, it cannot be traded or rented. There is no market for health but only markets for healthcare. In developing countries, especially the rural sector where technologies are labor-intensive, bad health leads to drops in current income and current productivity. If treatment is not sought or does not lead to a full recovery due to constraints in income or lack of health insurance, an individual will become more vulnerable to illnesses and injuries, likely result in drops in future productivity and future income. Utilizing the relationship between consumption and health illustrated in Figure 1.1, the dynamic impact of a poor health may be viewed as bending the budget constraint further toward the origin.

It is relatively easy to conceptualize this dynamic relationship, but empirical work in this area always encounters difficulties due to the limitations of data. Data on health, healthcare provision and health insurance from developing countries are very few, and for the ones that are available, health status are usually not recorded. To overcome the difficulties, some empirical research resort to measures such as the nutrients intake, caloric intake, Body Mass Index (BMI), days of ill, or medical records as proxies for health (see Ray, 1998, Deolalikar, 1988, Ye and Taylor, 1995).

One of the major empirical themes is to seek evidence in the relationship between nutrition and income. The null hypothesis under investigation is that income will result in better nutrition and vice versa. There is strong evidence in favor of the null at aggregate level, however, micro-level studies so far are not as supportive. For example, Behrman and Deolalikar (1987) use a special nutrition survey conducted by the ICRISAT during the 1976-77 and 1977-78 crop years in the ICRISAT India villages and find that the elasticity is actually close to zero.

Another major theme concerns the relationship between health and income fluctuations. As discussed earlier, the demand for health and other goals in life can be conceptualized as a conventional individual optimization problem. When economic environment changes, an individual will make adjustments along the indifference curve and incur necessary tradeoffs between health and the other goals as long as it is possible to fully access his expected income across all states of nature or over his lifetime. If he encounters either imperfect insurance or liquidity constraints, the tradeoffs between health and the other goals may end up as corner solutions.

Lacking access to formal financial institutions, households in developing countries sometimes may not always have full access to their (expected) disposable income, and will need to rely on their own wealth and social network to secure their wellbeing. One may wonder whether the tradeoffs between health and other commodities are optimal under the informal risk-coping arrangements. There have been some empirical efforts trying to analyze these tradeoffs. For example, when examining child health in rural South India, Behrman (1988) finds that because households are not able to smooth consumption, the health of children suffers during seasons before the major harvest, and girls are especially hard hit. Rose (1994) uses nationally representative data from India to show that the survival rate of girls relative to boys increases if income shocks are favorable in the early years of life, a result explained by discrimination coupled with the inability to smooth consumption. While investigating changes in the body size of children before and after major flooding in Bangladesh, Foster (1995) finds that body size suffers notably for household that were unable to borrow or insure against the associated income fluctuations. All empirical findings argue that health is sacrificed for rather short-term purposes, and long-term losses associated with such tradeoffs will curb economic growth. If we follow Morduch's method in calculating welfare losses, we should also take the long-term welfare losses into consideration as well.

1.3 TOWNSEND'S APPROACH

Abstracting from the actual smoothing strategies households may adopt, Townsend (1994) sets up a general equilibrium model and examines the effectiveness of mutual insurance in general with the data on three villages collected by the ICRISAT. His theoretical framework and estimation strategies are reviewed here since they are closely related to my research.

Townsend first quantifies the scope for insurance and justifies the grounds for choosing insurance rather than credits. According to his calculations, the coefficients of variation for many income sources are very high, suggesting high risks involved in these income sources and in overall income. In the next step, he derives the Pareto optimal conditions under full insurance in a multi-period static model. Individuals are assumed to have preference over consumption and leisure, with individual-specific characteristics such as age and sex embedded in the utility function. Assuming separability in consumption and leisure, he obtains a structural expression that relates the adult-equivalent consumption of an individual to that of the village (weighted) average. His later estimations are based on a reduced-form of the Pareto optimal conditions:

$$(1.6) \qquad c_t^{*j} = \alpha^j + \beta^j \bar{c}_t + \delta^j A_t^j + \varsigma^j X_t^j + u_t^j,$$

where c_t^{*j} is individual j's period t adult-equivalent grain consumption, $\bar{c}_t$ is village average grain consumption per adult equivalent, A_t^j is a demographic term summarizing

j's period t demographic feature and his different from that of the village weighted average, X_t^j is any other variable such as incomes from various sources, and u_t^j is a disturbance term. Assuming common coefficient of risk aversion and common welfare weights, full insurance implies that $\beta^j = \beta = 1$ and $\varsigma^j = \varsigma = 0$ for all j.

A nontrivial empirical problem that arises if (1.6) is taken directly into estimations is what to use for average consumption. The answer depends on whom we assume to be engaged in mutual insurance. One possibility is that the entire village is involved. In this case, we need the true village aggregate consumption to compute $\bar{c}_t$. However, only 40 households were surveyed in each village, and information on village aggregate consumption is missing. Townsend implicitly assumes the law of large numbers and uses the sample average as a proxy for the overall village average. However, this immediately runs into another problem: individual consumption in the sample must sum to the sample average, so that the average estimated coefficient on group consumption will be unity mathematically, and this may have nothing to do with insurance. To get around this problem, he runs a time-series regression for one household at a time over the sample period, and excludes the consumption of the specific household under examination. His estimates reject full insurance statistically.

To correct for measurement errors in explanatory variables, Townsend also estimates a difference equation as (1.7):

$$(1.7)\ \left(c_t^{*j} - \bar{c}_t\right) - \left(c_{t-1}^{*j} - \bar{c}_{t-1}\right) = \delta\left(A_t^j - A_{t-1}^j\right) + \varsigma_\Delta\left(X_t^j - X_{t-1}^j\right) + \left(u_t^j - u_{t-1}^j\right) - \varsigma_\Delta\left(v_t^j - v_{t-1}^j\right),$$

where the v's denote the i.i.d. measurement error with zero mean. (1.7) is estimated in a time series by household for each village. Now the test for full insurance is reduced to $\varsigma_\Delta = 0$. Estimates from this regression are not statistically different from the previous one.

Townsend also runs cross-sectional regressions for the landless and the landed households separately to see whether income terms are more likely to enter for the poor. In these regressions, the sample average consumption can be safely included in the right-hand side because it is no longer the average of the dependent variable. He finds that the landless seem to be more risk averse across all villages and are much less well insured against income shocks than the landed household in all villages.

1.4 MY RESEARCH

The important role health plays in determining current and future welfare in developing countries has been examined extensively in the literature of development economics and health economics. As with income shocks, health shocks incur direct losses in labor income because of losses of working hours and drops in productivity. Health shocks incur additional costs if medical treatment is sought and time is spent on seeking medical care and convalescence. Health shocks are also unique in the sense that there can be prolonged effects on future labor productivity, income, and welfare. As discussed earlier, there are other goals in life and good health is only one of them. Figure 1.1 illustrates the optimal solution in a static setting between health and consumption when there is no uncertainty, and the results are applicable in other static settings when health fluctuates with a known distribution and households have full access to their expected income under insurance.

In developing countries where access to formal credit or insurance markets is limited, protections from risks are managed through informal channels. One may naturally wonder whether the informal channels provide reasonably good insurance resembling the function of a formal health insurance. It is important to examine how individuals and households choose between consumption and health through the informal channels, so that public policies can be designed to incorporate local specifics without distorting private efforts.

The data set used to support this dissertation was a panel household survey collected by the International Crop Research Institute for Semi-Arid Tropics (ICRISAT) from three villages in southern India. It is the same the same data set used in Townsend (1994), Kochar (1995, 1999), Eswaran and Kotwal (1989), and many others in the literature of economic development.

This dissertation focuses on two key issues that have yet to be addressed in the literature. First, whether the informal channels can provide reasonably good insurance that resembles the functions of a formal health insurance. Second, when informal insurance is not perfect, which is mostly likely the case, whether health is forgone for the sake of current consumption, and who are less well protected against health risks. It takes a more general view of the theoretical implications of mutual insurance, discusses an empirical problem has been ignored by the existing literature using the ICRISAT data set, and tests for full insurance with several empirical and theoretical enhancements relative to Townsend (1994). The proceeding of this dissertation is as follows:

Chapter 2 describes the data set and the socioeconomic conditions of the ICRISAT sample villages. It provides summary statistics for key economic indicators such as in household incomes, consumption, demographic characteristics, and fluctuations in incomes.

Chapter 3 develops a static model of full insurance in which households are assumed to optimize over consumption, health, and leisure under uncertainties in health. No particular form of informal insurance is assumed explicitly. When dealing with health problems, individuals take sickness leave for recovery at the cost of labor income losses. Three types of specifications—constant absolute risk averse, constant relative risk averse, and nonseparable preferences are considered and theoretical implications compared in general equilibrium. Empirical estimation strategies are also discussed briefly.

The ICRISAT data set does not contain a direct measure of health or illness. Reports on taking sickness leave from work are used instead as a response variable toward health risks. When information on labor market participation is missing due to nonrandom reasons, imputation of sickness leave will lead to biases in estimations. Chapter 4 focuses on the problem of missing observations in the ICRISAT labor data. It explores the possible reasons for missing observations and the key factors determining an interviewee's observability in the village labor market. The results suggest that controlling for other factors, individuals from lower income households are more observable in the village labor market, which is strong evidence in support of intrahousehold risk sharing. This chapter then proceeds to discuss the biases in estimations of two existing empirical works using the same data set but without addressing the problem of missing observations. The results from this chapter are incorporated in the empirical estimations in Chapter 5 and Chapter 6.

Identifying data problems and methodological problems in Townsend (1994), Chapter 5 revisits Townsend's regressions with four more years of data, selection adjustments, and village-specific prices. After suitable corrections, the results suggest that mutual insurance is far less perfect as found in Townsend.

Based on the theoretical predictions from Chapter 3, Chapter 6 tests for full insurance with the full data set, village-specific prices, and bias adjustments developed in Chapter 4. Hypotheses on the separability between consumption and health are also tested along with full insurance. The results strongly reject full insurance as well as separability between consumption and health. It is shown that households tend to seek less sickness leave than optimum, and increase in sickness leave can improve welfare.

CHAPTER 2 THE ICRISAT DATA

This chapter offers a brief overview of the ICRISAT data used in empirical investigations in later chapters. Summary and analytical statistics are provided in the tables attached at the end of the chapter.

2.1 DATA COLLECTION

The data is from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Semi-arid tropical regions are often characterized by scanty and uncertain rainfall on which agricultural production largely depends, infertile soils, poor infrastructure, extreme poverty, rapid population growth, and high risks. The ICRISAT conducted a longitudinal household survey during 1975-1984 crop years with six villages from three semi-arid regions in India. Among the six study villages, only Aurepalle in Andhra Predesh, Shirapur and Kanzara in Maharashtra have complete ten-year data, therefore, discussions in this dissertation are specific to these three villages.

A fixed stratified sample of 40 household was randomly drawn from each village according to four landholding classes: landless, small farm, medium farm, and large farm; that is, 10 households from each class. Landless households were defined as those operating less than 0.5 acre land and whose main source of income was earnings in the casual agricultural labor market. Definitions for higher landholding classes were not uniform for different villages.

Data on household transactions, labor and draft power utilization, and crop cultivation by plot were canvassed at three-to-four week intervals. Information on the composition of the household, crop patterns by plot, debts and credit, stocks, livestock, implements and machinery, and farm building was updated annually. Table 2.1 has provided some summary statistics.

There are minor sampling problems as in other survey data sets. First, landless households are somewhat underrepresented in all the villages. On average across the six villages, they comprise about one-third of the households in the population of interest, but their share in the sample is only one-quarter. Nonetheless, it is considered as a fair reflection of their presence in the population of interest because their mean household size is less than that of landed households. Another issue is that the ICRISAT chose the villages that were willing to "cooperate". In order to provide the sample households with incentives to participate continuously in this free survey, the ICRISAT offered an annual three-to-four day excursion to places of agricultural and religious significance. Many of the participating households also realized that the biological research conducted on their and their neighbors' lands would benefit them the first if any technological components performed well. As a matter of fact, there was excess demand on the part of farmers to cooperate, and the investigators often chose farmers outside the panel to spread the potential benefits of the research activities across more households in the village. As mentioned in Chapter 1 of this dissertation, if one is interested in the risk sharing of the whole village while the true village aggregate variables are not available, problems of endogeneity, measurement errors, and small sample arise. Certain econometric techniques are then required for corrections.

2.2. THE STUDY HOUSEHOLDS

Agricultural Settings

Agriculture in the Semi-Arid Tropical (SAT) areas in southern India relies heavily on the rainfalls brought by the southwest monsoon. The monsoons usually span between mid-June and mid-October. However, the exact onset of the monsoon and the amount of rainfall to be received are highly uncertain. Nonetheless, planting in advance is necessary because delays are strongly associated with yield losses. When the monsoon initiation is erratic, the planned cropping strategy may no longer be optimal, and farmers will adapt to emerging information on rainfall events by changing crops or by fallowing land. In the formal terminologies of risks and insurance, we may well argue that farmers' planting decisions are dynamic and are *ex ante* efficient given the information available at the time. Farmers also engage in *ex post* efficient production activities based on the revealed information on monsoon initiation and percipitation. As shown in Table 2.1, Kanzara has steadier and higher amount of rainfall of all three study villages. Rainfall in Shirapur is the most scarce and scanty of all three study villages.

Household Demographic Composition

About one-half of the sample households are nuclear families; four-tenths are stem, in which one or both parents live with their son's family; and one-tenth are joint-stem, in which the parents reside with more than one married son. Despite of declining fertility, the population of the study villages is still growing. Seasonal, temporary migration to pursue employment opportunities is widespread and appears to be increasing (Walker and Ryan, 1990). For example, about 40% of the sample households in Shirapur and Kanzara ever contained a member who left the village.

Household Assets, Income, and Consumption

The value of land comprised up to 80% of total asset value in the study villages. There was no record of land sales during the survey years, which makes it feasible for us to take land as a fixed household factor or wealth. Productive assets include draft animals, farm building, and farm machines such as pumps and tractors. Consumer durable goods include residential buildings, household utensils, jewelry etc. They are individually less significant than landholding in the composition of household assets. Table 2.4 presents the relative ratio of adult-equivalent consumption in adult-equivalent household income for all landholding classes for each village. It can be viewed as the consumption share in total household income at per capita level. We can see that consumption share decreases as landholding class increases for all three villages, or equivalently, savings rate increases with landholding. Even though the average village-level consumption share is between 0.5 to 0.55, that is, average savings rate ranges between 0.5 to 0.45, landless households have significantly lower savings rates, ranging from 0.21 to 0.34.

As expected, food consisted of a very large proportion of consumption expenditure in the sample households. Food expenditure exceeded 60% of total expenditure for the vast majority of household in the three villages, especially in Aurepalle where mean income was the lowest.

There was negligible to mild inflation during the mid- to late 1970s across the three villages, rising inflation from 1979/80 to 1981/82, and consumer price stagnation onwards[1]. Using the 1980 prices, the mean real income was higher in Kanzara where agricultural growth has been more marked. The three-village mean per capita income was around Rs 700 during the survey years, which was considerably less than the all-India per capita income of Rs 1,080 for 1977. For regions impoverished as such, absolute

[1] All-India agricultural CPI from 1975 to 1984 were: 68.8, 75.1, 77.0, 91.0, 100, 104.3, 113.5, 128, 132.4, 139.5. Source: The World Bank and International Finance Corporation, International Economics Department Socio-Economic Data Division.

poverty may provide a more meaningful picture. Based on a consumption expenditure criterion, about three-fourths of the households in the study villages fell below the poverty line during the 10-year survey. Based on a net income criterion, the incidence of poverty was substantially less, ranging from 61% in Aurepalle to 43% in Kanzara, but nonetheless prevalent (Walker and Ryan 1990).

Crop revenue and labor earnings were the two dominant income sources in the study villages. The labor market in all three villages is large and active. Household wage income amounts to 25% of crop profits on average, although its importance varies by farm size. For landless households and small farm households, wage income exceeds crop income by 224%. Almost all households derived some income from the sales of livestock, but livestock production or animal trade was not as important as crop or labor income in any of the study villages.

Caste

Caste strongly influences many aspects of people's lives, and also contributes to explaining some more subtle dimensions of human behavior in the villages, for instance, the choice of occupations and diets. Most residents in the study villages are caste

Hindus. Relatively few of the caste Hindus are Brahmins. The majority belongs to communities whose traditional occupation is farming. There are at least a dozen of service-caste households residing in the study villages whose traditional occupations are carpenters, blacksmiths, barbers, water carrier, goldsmiths, etc. In terms of population, political power and wealth, there are one or two dominant castes in each village. However, caste does not play a direct role in the village agricultural labor market (also see Table 3.1).

Education

In the late 1970s, the mean literacy rates and educational levels of the sample households were not that different from the all-India average[2]. Also, like much of rural India, there is an increasing trend in education levels over time. Greater access to higher-level schooling within the study villages is one partial explanation, the other and more important part is that there is a growing awareness that a higher level of education is required for nonfarm employment and for a more modern, literate life.

Although some large farmers worry that increasing education will cause irreversible change to the agricultural labor force, most villagers consider education as a good thing for their children. Compared to their parents, children are more likely to enroll and spend more time in school. Boys are more educated than their fathers, and girls are more educated than their mothers. Education inequality by gender is quite prevalent in all study villages. In each age group, women are less educated than men. Even with the increasing trend of education, women in each age group progress much less. For example, the three-village education profile of women age fifteen to nineteen resembles that for men older than forty.

Infrastructure

All of the study villages are located no more than a few kilometers away from a paved road, and buses ply daily to and from the village or stop a few kilometers away. The villages are electrified, but relatively few of the households in each village receive electricity, which is primarily used for agricultural purposes. One fair-price shop selling subsidized kerosene, sugar, rice or wheat is located in each village. One cannot walk very far in any of the villages without encountering a small private shop retailing a few inexpensive consumer goods. The majority of villagers sell their produce in nearby regulated markets, which appear to be quite efficient and receive infrastructural support from the government. Sales rely on the auction system in the regulated market.

---

[2] According to World Bank (1980), about 64, 27 and 6 percent of individuals aged six to eleven, twelve to seventeen, and eighteen to twenty-three, respectively, attended school.

Health and Nutrition

In the late 1970s, sanitation was inadequate in all the study villages. Many household cooking areas were infested with flies and mosquitoes. Many households also had litter, debris, and / or animals within their living compound. Only about 15 percent of the sample households were considered to satisfy good conditions of household cleanliness. As for drinking water, coli-form bacteria were found in about 46% of the wells in the study villages, and the presence of E. Coli indicates fecal contamination, which may be responsible for the transmission of serious diseases.

None of the villages has a trained physician. Distance to the nearest government-operated primary health center ranged from 4 km to 13 km. The public health centers serving the study villages were poorly staffed and equipped. Although some were better stocked than others, frequently the only medicines available were aspirin, calcium lactate, and penicillin.

Because of these conditions, sickness is no stranger to the villages. Based on weekly interviews of 349 people in Dokur in 1982/83, the incidence of illness ranged from a weekly low of 14 percent to a weekly high of 28 percent during the fifty-two-week period of analysis (Bidinger, Nag, and Babu, 1986).

Another trait shared by residents of the six study villages was the dominance of cereals in supplying energy and nutrients in the diet. In most of the villages, the bulk of thiamin, iron, protein, niacin, riboflavin was derived from cereal consumption. The lack of diversity in the sources of some of the micronutrients such as riboflavin is usually not regarded as a good thing. For example, both children and adults in Aurepalle, where milk and milk products such as buttermilk loomed larger in the diet, had significantly higher riboflavin intakes than residents of the other villages.

Women's health and nutritional status are inferior to men's. Two material reasons are often put forward to explain the apparent discrimination against women: (1) they are of little social security value to their parents because they leave home after marriage and, (2) sons have better income-earning opportunities.

Focusing on children, more boys than girls of preschool (1 to 6 years) and school age (7 to 12) were poorly nourished on a standard of weight-for-age. Few of these children had second- or third-degree malnutrition on a weight-for-height index.

Seasonality of Nutrition and Health

Based on research and experience in tropical developing countries, weather and the associated agricultural activities tailor the seasonality of health and nutritional status. In particular, the wet season is purported to be the period when malnutrition and illness are at their peak in tropical developing countries. Malaria, diarrhea, and skin diseases are generally more prevalent at that time. Their effects are also more debilitating because the wet season is also when staple food supplies and intakes are dwindling and when labor demand and energy expenditure are rising.

The wet season in the six villages is generally from mid-June through September and it represents a period of lean food availability (Walker and Ryan, 1990). Other lean periods occur in the cold dry season in Shirapur because of the practice of monsoon fallowing and post-rainy season cropping and in the hot dry summer season in Aurepalle. With those two exceptions, the cold and dry season from November to February and the hot and dry season from March to May are times of seasonal food grain surplus. Contrary to expectations, Walker and Ryan (1990) find that the mean per capita intakes of energy and nutrients did not show a marked depression in the wet season. Indeed, more of the village by nutrient seasonal patterns fit an inverted U-shape than the expected dip in the wet season. In Aurepalle and Shirapur, nutrient intakes in the surplus food grain season were either equal to or in excess of the levels in the lean season.

Illness and Labor Market Participation

Using information on health and nutrition specially surveyed by the ICRISAT, Deolalikar (1986) finds evidence on the positive relationship between nutritional status and labor productivity in the study villages. His estimations show that the elasticities of the effect weight-for-height has on farm productivity range from 2.0 to 4.2. The mechanisms through which this relationship operates are still not clear. He also finds that men with superior anthropometric status select themselves or are selected for the more strenuous labor activities.

Lipton (1983) hypothesizes that overall labor participation rate in India is reduced 5 to 6 percent due to illness. The results from Walker and Ryan (1990) indicate that disabilities have more substantial effect on decisions to participate than Lipton suggests. They showed that men with physical disabilities who do participate in the labor force will do so for only 11 percent of the time compared to 56 percent for able-bodied participant males—a reduction of 45 percentage points. On the other hand, physically disabled females who participate in the labor force only reduce their participation from 58 percentage points to 55 percentage points, a reduction of 3 percentage points compared to able-bodied females.

One can imagine the large immediate benefits from improved health and nutrition to the study villagers, especially to men. As a matter of fact, if the difference is as high as suggested, it may provide a partial rationalization for men receiving first priority in the intrahousehold allocation of food. The frequent failure of special nutrition programs aimed at the nutritionally vulnerable groups, such as preschool children and pregnant and lactating women, may be a result of the diversion of the added food supplies to male members in order to generate increased wage income. If this is true, it implies that targeted nutrition programs may not succeed in the short term unless adult males are also included.

2.3 FLUCTUATIONS IN INCOME, CONSUMPTION AND THE DECISION TO TAKE SICKNESS LEAVE

Along with uncertain rainfall, factors such as commodity prices, crop yields, human sickness and injuries, diseases and death of plow animals also contribute to risks in agricultural productions and income. The data contains plot-level information on input and output. Input factors include labor and draft power utilization, pesticide, fertilizer, etc. Output factors include main and by-products. Continuous information on labor wages and commodity prices are also available. These can be used to compute income and profits from various sources in a time series.

One can obtain a measure of risks and gauge the scope of mutual insurance by constructing variance-covariance matrices using the income data. Coefficient of variation (CV) measures the riskiness of a random variable itself, whereas correlation measures the degree of covariation between two random variables. Townsend (1994) computes the CVs and correlation in key production variables such as crop mix, soil type, and crop-soil combination. He finds rather high CVs and low correlation across different groups, which indicates that diversification across crops will reduce risks. However, most households do not seem to hold a crop-soil portfolio. He argues that this may indicate the existence of alternative risk-reduction devices in the villages.

As stated earlier, the success in cropping depends on the timing and the precipitation of planting season rainfall. Figure 2.1 demonstrates monthly (cumulative) rainfall by village from 1975 to 1986. Monsoon rainfall was scantier and more erratic in Aurepalle and Shirapur, and higher and less variable in Kanzara. Upland and shallower soil generally cannot store enough moisture, hence require planting during the rainy season, whereas deeper soil retains moisture longer and it is much more profitable to cultivate crops during the post-rain season. Coupled with heterogeneous soil types and depths, managing risks in crop production is not an easy task.

In addition to uncertainties in weather patterns and crop failures, human illnesses and injuries are also prevalent among the sample households. Participating households were asked whether they have lost working hours because of illness or injury. The amount of time loss for each labor market participant can be inferred from a fairly detailed labor utilization data collected every three to four weeks during the 1975-84 survey years. Based on the kind of work and the payment schedule, income losses associated with illness and injuries can be computed. Information on medical spending, doctor fees and travel expenses in seeking medical treatments is recorded in a current consumption data, however, the observed values and frequencies are suspiciously low. Nonetheless, approximately 40 percent of the sample has reported loss of working hours during a typical survey year, and the highest number of such incidences goes as high as 18 times per year.

Table 2.1 Summary Statistics by Village During 1975-1984

Village Variable	Mean Aurepalle	Shirapur	Kanzara
Yearly Rainfall (*cm*)	625.16	650.70	851.31
	(196.24)	(217.18)	(186.29)
Landholding class[1] (*ha*)			
Landless	<0.2	<0.2	<0.2
Small farm	0.2-1.2	0.2-2.0	0.2-1.8
Medium farm	1.2-3.2	2.0-5.3	1.8-5.3
Large farm	>3.2	>5.3	>5.3
# of household members	6.70	7.23	7.17
	(2.60)	(2.97)	(3.60)
Age	52.20	49.42	44.87
	(12.40)	(10.57)	(11.08)
# of members in age 15-44	4.8	4.9	5.4
	(5.2)	(5.4)	(5.9)
# of members in age 45-64	1.7	1.5	1.7
	(3.5)	(3.4)	(3.9)
# of male members in age 15-44	2.6	2.5	3.6
	(4.6)	(4.4)	(5.1)
# of male members in age 45-64	1.1	0.9	0.6
	(3.1)	(2.8)	(2.5)
# of living siblings	3.22	3.66	3.38
	(2.07)	(2.19)	(2.14)
# of resident married sons	0.58	0.38	0.15
	(0.94)	(0.78)	(0.44)
Household Head			
Male percentage	24%	18%	24%
Age in 1980	33.0	30.0	31.0
	(19.0)	(18.4)	(18.4)
Percentage of at least primary education	42%	45%	52%
All income[2] (1980 Rs.)	586.45	539.26	778.81
	(617.40)	(393.31)	(720.35)
Labor income (1980 Rs.)	106.14	191.70	290.16
	(128.11)	(161.16)	(242.78)
Crop profit (1980 Rs.)	243.87	240.30	360.74
	(443.93)	(319.82)	(584.66)
Profit from trade and handicrafts (1980 Rs.)	123.76	18.12	23.51
	(238.78)	(145.26)	(93.31)
Profit from animal husbandry (1980 Rs.)	112.69	88.72	104.40
	(234.25)	(145.22)	(216.90)
All household consumption (1980 Rs.)	279.82	318.80	323.43
	(139.88)	(154.98)	(153.44)
Household grain consumption (1980 Rs.)	191.87	185.70	179.08
	(82.55)	(74.35)	(79.50)

Main occupation			
Farm percentage	34.24%	42.90%	34.57%
Nonfarm percentage [3]	65.76%	57.10%	65.43%
Education of non-head member			
Male of age 19+			
Female of age 19+			
Missing	10.82%	14.48%	16.74%
Illiterate	63.39%	41.43%	41.57%
Up to primary school	18.77%	5.36%	11.11%
Higher	7.02%	38.73%	30.58%

Note:
[1] Sources: Walker and Ryan (1990), Townsend (1994).
[2] Income is defined as monetary and imputed net returns to family-owned resources, encompassing family labor and owned bullocks, capital, and land.
[3] Including livestock rearing, trading, shop keeping, rural crafts, transportation, domestic, traditional caste occupations, school attending, etc.

Table 2.2 Coefficient of Variation and Correlation over Major Income Sources

Village	*Landholding*	*Crop Profit*	*Labor*	*Trade & Handicrafts*	*Animal Husbandry*
Aurepalle	Landless	2.87	0.72 0.06 (0.50)	1.64 0.33 (0.00) -0.25 (0.01)	2.18 0.17 (0.19) 0.47 (0.00) -0.18 (0.06)
	Small Farm	1.38	1.01 -0.22 (0.02)	1.23 0.69 (0.00) -0.35 (0.00)	1.79 0.35 (0.00) -0.02 (0.82) 0.18 (0.07)
	Medium Farm	1.39	1.53 -0.17 (0.07)	3.70 -0.10 (0.28) 0.05 (0.64)	1.53 0.32 (0.00) -0.07 (0.44) -0.05 (0.63)
	Large Farm	0.75	3.08 -0.26 (0.01)	2.33 -0.20 (0.04) -0.20 (0.04)	1.76 0.44 (0.00) -0.12 (0.22) -0.17 (0.08)
Shirapur	Landless	2.83	0.56 -0.21 (0.03)	3.08 -0.07 (0.45) 0.08 (0.41)	1.40 0.42 (0.00) -0.16 (0.09) 0.24 (0.01)
	Small Farm	1.49	0.99 -0.07 (0.46)	4.28 -0.11 (0.25) -0.11 (0.24)	1.30 0.22 (0.02) 0.26 (0.00) -0.05 (0.60)

Shirapur (cont.)	Medium Farm	1.20	0.82 0.35 (0.00)	4.02 -0.08 (0.39) 0.07 (0.48)	1.41 0.22 (0.01) 0.05 (0.63) -0.08 (0.41)
	Large Farm	1.00	1.54 0.00 (0.96)	3.89 -0.17 (0.09) -0.15 (0.14)	1.21 0.10 (0.32) -0.14 (0.15) 0.05 (0.58)
Kanzara	Landless	2.86	0.84 0.07 (0.50)	4.30 0.16 (0.09) -0.03 (0.76)	2.14 0.01 (0.88) 0.30 (0.00) 0.22 (0.02)
	Small Farm	1.38	0.56 -0.21 (0.03)	1.7 -0.11 (0.29) -0.15 (0.13)	1.82 0.49 (0.00) -0.17 (0.09) 0.12 (0.22)
	Medium Farm	0.88	0.90 -0.18 (0.06)	2.80 0.17 (0.08) -0.09 (0.34)	1.07 0.32 (0.00) 0.09 (0.38) -0.17 (0.08)
	Large Farm	0.89	1.45 0.12 (0.21)	5.76 -0.08 (0.41) -0.09 (0.36)	2.03 0.41 (0.00) 0.03 (0.74) -0.04 (0.68)

Note: Reported in parentheses are t-statistics.

Table 2.3a Coefficient of Variation and Correlation between Household Income and Major Consumption Components

Sample	*Landholding Class*	*Household Income*	*Total Consumption*	*Food Consumption*	*Grain Consumption*
All Villages by Landholding Class	Landless	0.54	0.56	0.54	0.65
			0.86	0.87	0.87
				0.91	0.90
					0.99
	Small Farm	0.47	0.44	0.42	0.58
			0.87	0.80	0.79
				0.74	0.70
					0.98
	Medium Farm	0.70	0.64	0.78	0.91
			0.83	0.77	0.81
				0.82	0.76
					0.96
	Large Farm	0.89	0.67	0.54	0.71
			0.61	0.75	0.70
				0.51	0.38
					0.91
All Sample		0.81	0.71	0.65	1.00
			0.70	0.79	0.80
				0.81	0.70
					0.91

Table 2.3b Coefficient of Variation and Correlation of Adult-Equivalent Income and Consumption

Sample	*Landholding Class*	*Adult-Equivalent Annual Income*	*Adult-Equivalent Annual Consumption*	*Adult-Equivalent Annual Food Consumption*	*Adult-Equivalent Annual Grain Consumption*
All Villages by Landholding Class	Landless	0.94	0.54	0.53	0.59
			0.49	0.66	0.82
				0.85	0.64
					0.88
	Small Farm	0.85	0.52	0.51	0.57
			0.55	0.69	0.82
				0.88	0.67
					0.89
	Medium Farm	0.92	0.57	0.59	0.77
			0.69	0.80	0.81
				0.84	0.72
					0.94
	Large Farm	0.97	0.74	0.67	0.88
			0.72	0.85	0.82
				0.72	0.60
					0.94
All Sample		0.96	0.65	0.59	0.73
			0.67	0.78	0.79
				0.77	0.61
					0.91

Table 2.4 Summary of Consumption Spending Shares by Landholding

Landholding Class	*Aurepalle*	*Shirapur*	*Kanzara*
Landless	0.743	0.655	0.788
	[0.203, 5.010]	[0.038, 1.299]	[0.145, 1.877]
Small Farm	0.572	0.515	0.678
	[0.111, 1.270]	[0.011, 1.638]	[0.0450, 2.061]
Medium Farm	0.451	0.619	0.579
	[0.029, 1.340]	[0.057, 1.804]	[0.034, 3.539]
Large Farm	0.340	0.433	0.338
	[0.033, 0.973]	[0.055, 1.191]	[0.028, 0.929]
Overall	0.502	0.549	0.551
	[0.029, 5.010]	[0.011, 1.804]	[0.028, 3.539]

Note:

1. Reported are the ratio of adult-equivalent consumption in adult-equivalent income for each household-year, or the consumption share in total household income for each adult-equivalent.
2. The minimum and maximum ratios are reported in brackets.

CHAPTER 3 CONSUMPTION, HEALTH, AND SICKNESS LEAVE IN GENERAL EQUILIBRIUM

3.1 INTRODUCTION

In developing countries, especially in the rural areas where technologies are labor-intensive, good health is of particular importance, because it not only translates into one's labor earnings and savings on medical care at present, it also affects one's health and productivity in the future. Compared with wealthier countries and regions where there is easy access to credits and insurance as protections against income and health risks, rural households in developing countries are much less fortunate. Due to problems of information flow, adverse selection and moral hazard in formal financial markets are severe. Coupled with weak enforcement mechanisms, formal financial markets often fail or do not function well for rural households in developing countries.

Theoretical and empirical literature in this area has focused on how these households cope with risks, and how effectively they are protected against risks in income and health. For example, Walker and Ryan (1990) documents that rural households from southern India surveyed by the ICRISAT engage in informal borrowing and lending through various types of social network. Rosenzweig and Wolpin (1993) find that households manage income risks by accumulating and depleting productive assets such as bullocks.

Ray (1998) documents that crop sharing and crop inventories are widely adopted by agricultural households in developing countries. Kochar (1995 and 1999) finds rural households smooth income and consumption by labor diversification. Rosenzweig and Stark (1989) show that migration and remittances from migrants are used to cope with household income risks. Theoretical research in this field focuses on the information flow and contractual nature of these informal risk-coping strategies, whereas empirical research examines the efficiencies and effectiveness of the informal risk-coping strategies in protecting households against various risks.

In this chapter, I develop a static general equilibrium model of community-wide full insurance based on Townsend (1994). A representative individual is assumed to derive utility from consumption, health, and leisure, and choose to take sickness leave when coping with idiosyncratic health risks. Equilibrium conditions for consumption, leisure, and sickness leave are derived under three different assumption about preferences: constant absolute risk averse (CARA), constant relative risk averse (CRRA), and nonseparable preferences in health and consumption. Under CARA and CRRA preferences, individual consumption, sickness leave, and leisure can be expressed in terms of weighted village average consumption, sickness leave, and leisure, and a set of individual-specific factors. In the simplest case where the population is homogenous, theoretical predictions under both preference specifications suggest that individual choice variables is equal to the corresponding village average choice variables and nothing else. It is argued here that when testing full insurance in reduced forms, we are not able to identify which type of preference the study individuals actually have. In the third preference specification where consumption and health are assumed to be nonseparable, linear approximations are used to derive equilibrium consumption, sickness leave, and leisure. It is shown that when the elasticity of substitution between consumption and health is positive, i.e., the marginal utility of consumption increases with small increases in health and vice versa, consumption and sickness leave in equilibrium adjust according to a varying negative ratio across states. Estimation strategies are discussed under each of the three preference specifications.

3.2. THREE PREFERENCE SPECIFICATIONS AND GENERAL EQUILIBRIUM

The models presented in the following are based on Townsend (1994) with the extension of health as an objective in preferences, and individuals optimally taking time off from work when afflicted with health problems. Theoretical implications are compared with that of Townsend's. Townsend's is a static model of full insurance in multiple periods in which a set of idiosyncratic risks realize at the beginning of each period, and the realization is completely independent of past history. The models introduced here are all one-period full insurance general equilibrium models for notational ease.

Let us start by considering a closed village community with a heterogeneous population of N individuals, who are different in age, sex, welfare weights, and risk attitudes. As Townsend (1994) has pointed out, there are *a priori* grounds for taking villages as the natural unity of study because village economies satisfy "...the explicit and implicit conditions of general equilibrium". For example, many families have been present for generations; villages usually have their own legal systems replete with contract enforcement mechanisms; and village residents may have relatively good information about the ability, effort, and outputs of one another.

A representative individual derives utility from consumption C, health H, and leisure L. Each individual is assumed to have 1 unit of time endowment and allocates it across work, leisure, and sickness leave for recovery when afflicted by unexpected health problems. Denote $\overline{H}_i$ as individual i's health endowment, which is i's natural health status if he or she is not afflicted with illness or injuries. $\overline{H}_i$ is the upper boundary of i's health status. Assume that $\overline{H}_i$ is exogenously determined. Suppose there are total S possible states of nature in which individual i receives an i.i.d. health shock ε_{is}, which will reduce his or her health status from $\overline{H}_i$ if no measure is taken to seek recovery. Suppose that the number of individuals alive and residing in the village is large than or equal to the dimensions of health shocks realized in any state, so that it is feasible to smooth health risks at village level through mutual insurance if there is no market failure. Also, as $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N}\varepsilon_{is} \to \bar{\varepsilon}$, where $\bar{\varepsilon}$ is the mean of health shocks. The i.i.d. assumption on ε_{is} excludes all possible covariate health problems resulting from bad sanitation or contagion.

When afflicted with health problems, individual i chooses to take sickness leave T_{is} for recoveries. In particular, individual i's health status in state s is given by

$$(3.1) \quad H_{is} = \overline{H}_i + \sigma_i (T_{is} - \varepsilon_{is}),$$

where H_{is} is i's realized health status in state s, and σ_i has the interpretation of the rate of recuperation that varies across individuals. For example, young adults usually recover from diseases more quickly than seniors, therefore, when targeting the same level of health status, young adults can choose shorter sickness leave than seniors. Suppose that there exists a health threshold under which the individual will die, and let the threshold be normalized to 0. Assume that $0<\sigma_i \leq 1$ for those who are alive and present in the village, and let $\sigma_i = 0$ if death is triggered. As the recuperation rate gets closer to zero, recoveries require longer and longer sickness leave, and when death is triggered, nothing can revoke it.

It can be shown easily that when there is no health risk, $\varepsilon_{is} = 0$, sickness leave will not be sought in equilibrium because it incurs losses in labor income and it does not improve i's health status to a level higher than the natural upper bound $\overline{H}_i$.

Health can be viewed as a static version of a stock variable in (3.1). Adopting the conventional terminology for stock variables, health "depreciates" at rate σ_i because of an independent and idiosyncratic shock ε_{is}. Sickness leave, T_{is}, serves as investment and its effectiveness also depends on σ_i. The difference between sickness leave and health risks, $T_{is} - \varepsilon_{is}$, has the interpretation of net investment in health. Sickness leave and health risks are formulated in a linear fashion in (3.1). In other words, it is assumed here that the rate of return to net investment in health is constant, unlike the more comprehensive view in the literature of health economics that there is decreasing return to investment in health (see Grossman 1972, 2000, and Zweifel and Breyer, 1997). The linear formulation is adopted for the ease of derivations and the consistency in the models introduced in this chapter without loss of generality.

As mentioned earlier, the ICRISAT data set does not contain direct information on health or health risks. The only variable contains some information on health and health risk is the reported incidences of individuals taking time off from work because of sickness. Upon each interview, an individual would be asked whether or not he or she had lost work since the last visit. If the answer was yes, the investigator would further ask about the reasons for such losses—whether it was due to sickness, or involuntary unemployment, or religious and social events, etc. In principle, it is possible to impute the actual amount of sickness leave based on the information on labor utilization. However, hours of leisure and home production were not recorded, which has made such imputations infeasible. A problem is that sickness leave is treated as a continuous variable in (3.1), however, it is actually discrete in the data—either binary (= 1 if had taken sickness leave since the last visit, = 0 if not) or counts (= M times during the span of N interviews of a year). Chapter 4 and Chapter 6 of this dissertation will have further discussions on this problem.

Treating health as a stock variable has a long history in the literature of theoretical health economics. One of the most influential works is by Grossman (1972). Health in the Grossman model is treated as a stock variable that depreciates over time and recuperates by seeking medical treatment and spending time for convalescence. Later works based on the Grossman model have revealed that the real economic costs incurred by health risks include all monetary expenditures and opportunity costs associated with seeking treatment—monetary expenditures such as health insurance premiums, doctor fees, payments to medication and operations, travel expenses, and an opportunity cost, the losses in labor income due to time spent on seeking treatment and convalescence. The monetary side of health investment, however, is excluded from (3.1) to incorporate the fact that rural households in developing countries usually do not have access to formal health care or health insurance, and the only cost incurred here is the opportunity cost resulting from taking sickness leave.

3.2.1 SEPARABILITY IN PREFERENCES AND PREFERENCE SHOCKS

Separability in Preferences

When assuming certain preference specifications, one implicitly imposes restrictions on the equilibrium solutions sets. Theoretical implications are largely driven by the underlying assumptions on preferences. It is always a challenge to establish a theoretical model that is as explicit and tractable as possible, yet without loss of generality. Therefore, before jumping into the details of model specifications, it is important consider the nature of an individual's preferences over consumption, health, and leisure.

When considering a broad assortment of commodities, it is natural to partition them into separate groups, so that the preference ordering of one commodity group is independent of the other commodity groups. For each partition, there may exist subgroups of commodities that are mutually independent, as if there exists a utility tree (Figure 3.1). Accordingly, utility functions can be assumed to have special functional forms (see Technical Appendix A). For example, if preferences over all P groups of commodities are independent and weakly separable, the corresponding utility function has the form of

$$(3.2) \quad u(\mathrm{x}) = U(u_1(x_1), u_2(x_2), \ldots, u_P(x_P)).$$

In our case, if a villager's preference ordering on consumption goods does not change with respect to changes in his or her health status, all consumption items can be grouped together as one consumption aggregate separately from health. According, the utility function for this individual can be written as $u(C,H)=U(\phi(C),\varphi(H))$, where C represents the consumption aggregate, and H represents health status.

Stronger structural forms are obtained by introducing additional independence relations. For example, if preference orderings are weakly independent with respect to the P groups of commodities and with respect to the unions of all subsets within the commodity space, we have strongly independent preferences, which implies a strongly separable utility specification such as

$$(3.3) \quad u(x)=U\left(\sum_{i=1}^{P} u_i(x_i)\right)$$

A very special case of strong separability of utility functions is an additive utility function

$$(3.4) \quad u(x)=\sum_{i=1}^{P} u_i(x_i).$$

In our case, the utility function is given by $u(C,H)=\phi(C)+\varphi(H)$ if preferences are strongly independent.

It is relatively easy to conceptualize the independence of preferences theoretically. When dealing with real-life cases, however, it may not be as clear-cut and straightforward. Theoretical formulations cannot exhaust all of the possible commodities in the commodity space, therefore, we need to draw a fine line as for what shall be included or excluded in the utility functions. On the other hand, the exclusion of certain commodities may cause biases and misinterpretations in empirical estimations based on the theories.

For example, utility functions are customarily assumed to have only a few arguments, such as consumption, health, leisure, and sometimes wealth or income. In most of the literature of consumption smoothing and permanent income hypothesis (PIH), utility functions usually have only one argument—consumption. The implicit assumption is that all consumption items are strongly independent from other commodity partitions, so that we can group all consumption items into a consumption aggregate. Utility maximization under strong independence in preferences (additive utility functions) can be carried out by maximizing the utility for each independent partition. PIH for this one-argument utility specification predicts that if the capital markets are perfect, an individual's consumption depends on his or her expected lifetime income and nothing else. The welfare implication is that Pareto optimum is achieved with perfect capital markets. If one interprets the theory mechanically, testing consumption smoothing and the PIH is equivalent to testing whether or not consumption is excessively sensitive with respect to factors other than permanent income, such as transitory (cyclical) incomes, wage rates, interest rates, and wealth. This gives the rise of the decades-old yet still on-going debate as for how we measure consumption (market goods? home goods?), what we mean by excess sensitivity, how to interpret the excess responses, and how to evaluate the efficiency of the capital markets etc. (see, for example, Baxter and Jermann, 1999).

The core of the controversies lies in the interpretation of Pareto optimality and the choice of the dependent variable. The essence of Pareto optimality is that when capital markets are perfect, the welfare (utility) of an individual is smoothed throughout time, regardless transitory changes. In other words, it implies welfare smoothing throughout time when capital markets are perfect. In the literature of PIH, consumption is chosen as a metaphor or proxy for welfare under the implicit assumption that it is independent from other commodities in the commodity space. Following the same methodology, testing PIH and Pareto optimality can be extended to testing the smoothness of the path of any independent partition (such as leisure) within the commodity space.

Only when the chosen proxy (the dependent variable) is indeed independent in preferences, can the smoothness of its path be viewed as evidence for Pareto optimum. The difficulty is that we do not know whether or not real preferences are independent on the proxy. If the proxy is not independent from other commodities in preferences, it is still possible to find "excess sensitivity" when the capital markets are perfect, because factors determining other partitions that are nonseparable from the chosen proxy will affect the proxy itself. For example, Baxter and Jermann (1999) find that allowing for nonseparability between market goods and home goods in preferences, the "excess sensitivity" in market goods is actually compatible with the PIH. In other words, when treating market goods as an independent partition in preferences, the consumption path of market goods alone may seen to be excessively sensitive to factors such as cyclical wage rates and prices, which leads to the rejections of the PIH. If one stands one step higher in the utility tree (see Figure 3.1) and considers the nonseparable composite of market goods and home goods, testing the PIH is actually testing whether or not the composite is excessively sensitive. The "excess sensitivity" of market goods alone shall not be taken entirely as evidence against the PIH. To sum up, the correct test for the PIH relies on the correct choice of the dependent variable.

Nonseparability in preferences has been applied to business cycle macroeconomics. In micro-level household studies, however, its importance is under-emphasized. For example, in the literature of informal mutual insurance (see discussions in Chapter 1 of this dissertation), testing full insurance and Pareto Optimality is usually carried out by testing whether or not factors such as incomes, wage rates, or family structures affect the consumption of an individual. If there is perfect insurance, the consumption of an individual is explained by the community-wide aggregate consumption and nothing else. The problem of this literature is that it has taken the theoretical implications of Pareto optimality too literally. Similar to the PIH, Pareto optimality implies that welfare is protected against idiosyncratic risks. If we are able to partition the commodity space into mutually independent groups of commodities, Pareto optimality implies that each of the independent groups are protected against idiosyncratic risks. As in the PIH, the empirical challenge here is how to identify the independent partitions and how to construct the correct dependent variables for the tests.

In Townsend (1994), for example, utility function is additive in consumption and leisure. Other important factors such as health are excluded from the theoretical discussions and only come back in empirical estimations. In other words, the implicit assumption is that consumption and leisure are strongly independent from each other in preferences, and are strongly independent from other factors in the commodity space. From the previous discussions, we know that there is no need to include leisure in the utility specification at all, since it is strongly independent from all other commodities in preferences. Consumption alone can be used as the dependent variable to test full insurance. Then the natural questions are—Is consumption indeed independent from other commodities in preferences? If yes, how to measure consumption (e.g., market goods and/or home products)? If not, what are its nonseparable partitions? How to test full insurance properly?

Health and consumption are conventionally modeled to be strongly separable in utility specifications, even in the literature of health economics. However, are they indeed independent in preferences? Can higher consumption translate into better health and vice versa? Quite a number of empirical works have documented the nutrition link between health and consumption. The consensus is that higher food consumption provides better nutrient and caloric intake, hence improves one's health. However, this link seems to disappear if the income of an individual exceeds a certain level (see Ray, 1998, and Dasgupta and Ray, 1990). A more pertinent question, perhaps, is whether or not it is proper to assume independence between health and consumption for the individuals under study. It is more appropriate to assume independence between health and consumption for individuals with higher incomes than for those who live close to substance level. Accordingly, when testing full insurance for higher income individuals, it is proper to use consumption alone as the dependent variable. If consumption alone is excessively sensitive toward factors other than the village aggregate consumption, one can reject full insurance for these households. When testing full insurance for very low-income individuals, the dependent variable shall be the composite commodity of health and consumption together. The sensitivity of consumption alone or health alone does not imply that insurance is imperfect. The sensitivity of the whole composite is the key.

In the following, three types of preferences are assumed and theoretical predictions are discussed. The first specification is an additive constant absolute risk-averse utility (CARA) that corresponds to strong independence in preferences over consumption, health, and leisure. The second specification is a multiplicative constant relative risk-averse utility (CRRA), which corresponds to the weak independence in preferences. The third specification is nonseparable in consumption and health, but additively separable in leisure, i.e., leisure and the consumption-health composite are independent from each other in preferences.

Figure 3.2 illustrates the utility trees for these three specifications. Both additive and multiplicative utilities imply that health is independent from consumption in preferences, therefore, they share the same utility tree structure (see the top panel). When health and consumption are not independent in preferences, but nonetheless independent from leisure in preferences, consumption and health form a nonseparable composite. Together they share a knot in the utility tree that is parallel to leisure.

Preference Shocks

Another issue associated with independence in preferences is preference shocks. In some of the literature of risk sharing and consumption smoothing, preference shocks are introduced in utility specifications. The utility specifications either involve only one argument—consumption, or involve multiple yet additively separable commodities (see Mace, 1991, and Cochrane, 1991). From the previous discussions, we know that a single-argument utility specification and an additively separable multiple-commodity utility specification share the same implicit assumption that preferences are strongly independent, and yield the similar theoretical implications about complete risk-sharing. They are essentially the same.

The essence of including preference shocks in strongly independent preference (though not discussed in most papers) is to account for unobserved changes in other commodities in preferences that are not independent in preferences. For example, when consumption and health are not independent in preferences, if one adopts a single-argument utility function for consumption alone, the "preference shocks" are in fact unobserved changes in health, which have nothing to do with the efficiency of the underlying capital markets or insurance arrangements. If we allow for nonseparability in consumption and health and consider the composite as a whole, the "preference shocks" to consumption is no longer a shock to preference, but rather, a change in the substitution pattern between consumption and health due to nonseparability.

Part of the literature acknowledges the existence of "preference shocks" and tries to peel them off in order to fully examine the degree of risk-sharing. For example, when testing risk sharing, Mace (1991) introduced a single-argument utility function in which utility is derived from consumption and affected by a preference shock. The source and nature of this preference shock is not discussed. She considered both idiosyncratic and aggregate shocks to incomes in addition to the "preference shocks". Cochrane (1991) had a similar formulation when testing the PIH. In contrast, some researchers simply restrict their attentions to a world where there is no "preference shocks". For example, Eichenbaum, et al (1988) used a "no-preference-shock restriction" to identify and estimate the parameters of preferences and the household technology. To sum up, preference shock formulations can be viewed as a reduced-form fashion of incorporating nonseparability in preferences. If we formally model nonseparability, preference shocks are no longer an issue.

3.2.2 CARA PREFERENCES

In this section, individuals are assumed to have constant relative risk-averse (CARA) preferences over consumption, health, and leisure. Preferences are assumed to be additively separable (strong independence) as in Townsend (1994). Additive separability imposes very strong restrictions on the substitution patterns between consumption, health, and leisure. In particular, the elasticities of substitution[1] across all utility arguments are equal to zero. It suggests that utility maximization is equivalent to maximizing each separable utility independently then taking the summation. Individual i's utility in state s is given by

$$(3.5) \quad U^i(C_{is})+V^i(L_{is})+W^i(H_{is}).$$

It is assumed implicitly that a better health status does affect the marginal utility of consumption or leisure and vice versa, i.e., the elasticities of substitution across all utility arguments are zero. It is shown in the following derivations that consumption, health, and leisure interact only through the aggregate budget constraint. Formulations such as this have very strong theoretical and empirical appeals because they provide well-behaved equilibrium conditions that can be estimated structurally.

[1] If the utility function is given by $U(X,Y)$, the elasticity of substitution between X and Y is given by $\xi_{XY} = -\frac{d\ln Y/X}{d\ln U_X/U_Y}$.

Equilibrium Conditions

In equilibrium, the marginal rate of substitution between any two of the utility arguments is equalized:

$$(3.6)\quad \frac{\partial U}{\partial C_{is}} = \frac{\partial V}{w_i \partial L_{is}} = \frac{\sigma_i \partial W}{w_i \partial H_{is}}.$$

Following Townsend (1994), exponential utility functions for the additively separable preferences are assumed as follows:

$$(3.7)\quad U^i(C_{is}) = -\frac{1}{\alpha}\exp(-\alpha C_{is}),$$

$$(3.8)\quad V^i(L_{is}) = -\frac{1}{\beta}\exp(-\beta L_{is}), \text{ and}$$

$$(3.9)\quad W^i(H_{is}) = -\frac{1}{\gamma}\exp(-\gamma H_{is}).$$

Remember that the only source of risks considered in the model is health risks, which reduce health status, affects leisure and labor income time if sickness leave is taken. From this point of view, all utility commodities—consumption, leisure and health—involve randomness, directly or indirectly resulting from health risks, and the parameters α, β and γ reflect the degrees of absolute risk averse in preferences. It is assumed that all individuals have the same risk attitude for the ease of computation.

Now the programming problem for the determination of Pareto-optimal allocations can be written. Let λ_i denote the welfare weight associated with individual *i* that satisfies

$$(3.10)\quad 0<\lambda_i<1,\ \sum_{i=1}^{N}\lambda_i=1.$$

Let p_s denote the probability that state s is likely to take place. $\sum_{s=1}^{S}p_s=1$. Once a state is realized, individuals' health risks are determined $(\varepsilon_{1s},\varepsilon_{2s},\ldots,\varepsilon_{Ns})$.

Pareto-optimal allocations are solved by

$$(3.11)\quad \underset{\{C_{is},L_{is},H_{is}\}}{Max}\ \sum_{i=1}^{N}\lambda_i\sum_{s=1}^{S}prob(s)(U_{is}+V_{is}+W_{is}),$$

subject to resource constraints

$$(3.12)\quad \sum_{i=1}^{N}C_{is}\le\sum_{i=1}^{N}w_{is}(1-L_{is}-T_{is}),\ \text{for all } s,$$

where w_{is} denotes the wage rate realized in state s for individual i,

$$(3.13)\quad 0\le L_{is}\le 1,\ 0\le T_{is}\le 1,\ 0\le L_{is}+T_{is}\le 1,$$

and nonnegativity constraint

$$(3.14)\quad C_{is}\ge 0, L_{is}\ge 0, H_{is}\ge 0.$$

The resource constraints state that village aggregate consumption is no higher than the village total labor income which is the only source of income in any state, and total labor income depends on individual wage rates as well as work time. Health shocks directly affect individual's wellbeing by reducing utility from health. Sickness leave offsets these shocks, but incurs income losses by reducing total working time.

Let μ_s denote the Lagrange multiplier on (3.12) in state s. If nonnegativity constraint (3.14) on consumption and the resource constraint (3.13) on leisure and sickness leave are not binding, the first-order conditions for individuals present and alive on consumption, leisure and sickness leave are:

(3.15) $\lambda_i \exp(-\alpha C_{is}) = \lambda_j \exp(-\alpha C_{js}) = \mu_s$ for all i, j, and s,

(3.16) $\frac{\lambda_i}{w_{is}} \exp(-\beta L_{is}) = \frac{\lambda_j}{w_{js}} \exp(-\beta L_{js}) = \mu_s$ for all i, j, and s, and

(3.17) $\frac{\lambda_i \sigma_i}{w_{is}} \exp(-\gamma(\overline{H}_i + \sigma_i(T_{is} - \varepsilon_{is}))) = \frac{\lambda_j \sigma_j}{w_{js}} \exp(-\gamma(\overline{H}_j + \sigma_j(T_{js} - \varepsilon_{js}))) = \mu_s$

for all i, j, and s.

The common term p_s that should appear on the left-hand side of (3.15), (3.16) and (3.17) has been factored out and placed in the Lagrange multiplier. Equations (3.15)-(3.17) imply that the marginal utilities of consumption, health, and leisure are equalized across all individuals in the village for any state s, and the marginal utilities of consumption, health, and leisure for an individual is equalized across all possible states. These are the standard results of full insurance at the village level as well as individual level.

The utility function is formulated in a fashion that individual demand for consumption, leisure and sickness leave can be expressed in terms of the village aggregates or averages and individual characteristics. Let us denote village aggregate consumption in state s as

$C_s = \sum_{i=1}^{N} C_{is}$, aggregate leisure as $L_s = \sum_{i=1}^{N} L_{is}$, and aggregate sickness leave as $T_s = \sum_{i=1}^{N} T_{is}$. Summing up (3.15)-(3.17) across all individuals, we can derive individual i's demand for consumption, leisure and sickness leave in state s in terms of the deviation of individual characteristics from that of the weighted village average characteristics and village aggregates:

(3.18) $$C_{is} = \frac{1}{\alpha} \times \left(\log \lambda_i - \frac{\sum_{j=1}^{N} \log \lambda_j}{N} \right) + \frac{C_s}{N} \text{ for all } i, j, \text{ and } s,$$

(3.19) $$L_{is} = \frac{1}{\beta} \left(\log \lambda_i - \frac{\sum_{j=1}^{N} \log \lambda_j}{N} \right) + \frac{1}{\beta} \left(\frac{\sum_{j=1}^{N} \log w_{js}}{N} - \log w_{is} \right) + \frac{L_s}{N} \text{ for all } i, j, \text{ and } s,$$

(3.20)

$$T_{is} = \frac{1}{\gamma} \times \left(\frac{\log \lambda_i}{\sigma_i} - \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \frac{\log \lambda_j}{\sigma_j} \right) + \frac{1}{\gamma} \times \left(\frac{\log \sigma_i}{\sigma_i} - \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \frac{\log \sigma_j}{\sigma_j} \right) + \left(\frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \right) \times T_s$$

$$+ \frac{1}{\gamma} \times \left(\frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \frac{\log w_{js}}{\sigma_j} - \frac{\log w_{is}}{\sigma_i} \right) + \left(\frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \frac{\overline{H}_j}{\sigma_j} - \frac{\overline{H}_i}{\sigma_i} \right) + \left(\varepsilon_{is} - \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \varepsilon_{js} \right),$$

for all i, j, and s,

Welfare weights appear in the first term of equations (3.18)-(3.20). When factors such as the degrees of risk averse, recuperation rate, natural health status, and the severity of health risks are the same, equations (3.18)-(3.20) are reduced to the deviation of individual welfare weights from that of the weighted village average and an even share of the village aggregate. These are the typical results from complete risk-sharing. An individual with a welfare weight higher than the weighted village average welfare weight receives the even share of village aggregate consumption or leisure or sickness leave, plus a "premium" that is positively correlated with the net deviation. In other words, the higher the welfare weight of an individual, the higher that individual's consumption, leisure, and sickness leave will be in equilibrium.

Wage rate does not explicitly enter in (3.18) where individual consumption is determined, but it does not mean that wages do not affect consumption at all. Wages help determine the village aggregate resources, and the actual resource distributions are carried out based on individual characteristics. When controlling for the degree of risk-averse, welfare weights, and village aggregate income, equation (3.18) implies that an individual consumes an equal share of the village aggregate consumption as any other villager regardless of his or her wage rate.

Wage rate explicitly enters in (3.19) and (3.20). Its role is similar to that of welfare weights in that an individual with a wage rate higher than the weighted village average takes shorter leisure and short sickness leave in equilibrium, and the "discount" is proportional to the deviation of individual wage rate from that of the weighted village average.

Because of the additive separability in consumption, health, and leisure in preferences, the rate of recuperation σ_i does not enter (3.18) or (3.19) where individual demand for consumption and leisure is determined. It only appears in (3.20) where the demand for sickness leave is determined. To understand its impact on sickness leave and health, let us assume all other individual characteristics are constants, so that (3.20) is reduced to

$$(3.20)' \qquad T_{is} = \frac{1}{\gamma} \times \left(\frac{\log \sigma_i}{\sigma_i} - \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \times \sum_{j=1}^{N} \frac{\log \sigma_j}{\sigma_j} \right) + \left(\frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}} \right) \times \left(T_s - \sum_{j=1}^{N} \varepsilon_{js} \right) + \varepsilon_{is}.$$

This expression implies that an individual with a recuperation rate higher than the weighted village average will take a smaller share of the total net investment in health because the ratio $\frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}}$ decreases as σ_i increases. Conditioning on σ_i, an individual will take longer sickness leave when facing more severe health problems because ε_{is} enters additively. Combining (3.1) and (3.20)', we know that $H_{is} = \frac{T_s - \gamma \times \sum_{j=1}^{N} \varepsilon_{js}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}}$.

When the village is large enough, $N \to \infty$, we have $H_{is} \to \frac{T_s - \gamma\bar{\varepsilon}}{\sum_{j=1}^{N} \frac{1}{\sigma_j}}$, a constant for all individuals, that is, health status is smoothed for all individuals in each state s.

Controlling for other individual-specific factors, an individual with a higher initial health status than that of the weighted village average takes shorter sickness leave, and an individual encountering a larger health risk relative to that of the weighted village average takes longer sickness leave in equilibrium. This result resembles that of perfect health insurance in which total expenses associated with health risks are fully covered, so that an individual will take necessary and sufficient amount of economic resources for recoveries. Unlike a health insurance arrangement, (3.20) does not involve a reimbursement scheme explicitly. The reimbursement can be viewed as individuals' full access to the village aggregate resources implied by the resource constraints. It is the community all together sharing the expenses that individuals incur.

If the population is homogeneous, (3.18) and (3.19) are reduced to $C_{is} = \frac{1}{N} C_s$ and $L_{is} = \frac{L_s}{N}$, implying that all individuals behave the same regardless of the realization of the idiosyncratic health risks. Following the same assumptions, (3.20) is reduced to

$$(3.21) \qquad T_{is} = \frac{T_s}{N} + \varepsilon_{is} - \bar{\varepsilon},$$

that is, individuals facing larger health risks will take longer sickness leave. Plug in this expression back to (3.1) where health status is determined, we can see that all individuals eventually have the same health status along the optimal path, $H_{is} = \overline{H} + \sigma\left(\frac{T_s}{N} - \overline{\varepsilon}\right)$. We can see that health risks affect consumption and leisure only at the aggregate level in equilibrium.

In summary, under CARA and additive separable preferences, the relationships between health and consumption, and health and leisure solely depend on the income effect under full insurance, and there is no substitution effect because health is assumed to be separable from consumption and leisure.

The assumption of additive separability may be too strong because there are cases in which health is not additively separable from consumption, i.e., the elasticity of substitution between consumption and health is not zero. For example, good health is usually viewed as an unalienable part of enjoying high consumption, and consumption is also found to affect health and productivity through nutrition especially for very low 4income households (see Ray, 1998, and Zweifel and Breyer, 1997). The first case is compatible with Cobb-Douglas preferences. Under Cobb-Douglas specifications, the marginal utility of consumption depends on health and vice versa. Consumption and health adjust according to a constant rate in equilibrium (the elasticity of substitution between consumption and health is equal to –1 for all income levels). The second case involves a nonlinear relationship between consumption and health, that is, the elasticity of substitution between consumption and health is no longer a constant for all income levels. We need a general nonseparable preference specification in this case. Important theoretical and empirical questions would naturally arise: Will the theoretical implications on full insurance be the same for all preference specifications? How can we test for full insurance empirically in each preference regime? Are there flexible tests for full insurance that are independent of preference specifications? The following sections focus on the theoretical implications and empirical estimation strategies under these two regimes of preference specifications. Section 3.2.2 addresses the first case in which consumption and health are kept at a constant ratio in equilibrium. Section 3.2.3 addresses the second case where the elasticity of substitution between consumption and health is not a constant.

3.2.3 CRRA PREFERENCES

In this section a constant relative risk averse (CRRA) specification is applied to preference. To overcome the limitations of additive separability in consumption and health and to obtain theoretical implication in a more flexible framework, it is assumed in this section that consumption and health enter into the utility function in a multiplicatively separable fashion. Since leisure is not the focus of this dissertation, it is assumed to be additively separable as in Section 3.2.2.

In particular, let us consider a utility function such as $U(C_{is}, H_{is}) + V(L_{is})$, where $U(C_{is}, H_{is}) = \frac{(C_{is}^{\gamma} H_{is}^{1-\gamma} - 1)^{1-\alpha}}{1-\alpha}$ (see Picone et al, 1998, and Dardanoni, 1988), and $V(L_{is}) = \frac{L_{is}^{1-\beta}}{1-\beta}$. α and β are the coefficients of relative risk aversion, and γ has the interpretation of conventional Cobb-Douglas utilities where the marginal rate of substitution between consumption and health is equal to $\frac{\gamma}{1-\gamma}$. Assume $\alpha \in (0,1)$, $\beta \in (0,1)$, and $\gamma \in (0,1)$. The Arrow-Pratt coefficient of relative risk aversion with respect to health is equal to $\gamma - \alpha(\gamma - 1)^2$. The elasticity of substitution between consumption and health is equal to –1, suggesting that consumption and health are still

---

[2] Dardanoni (1988) derives the formula for the Arrow-Pratt coefficient of relative risk aversion for a two-argument utility function. The formula is: $R_H^R = -H \times U_{HH} / U_H$, which is the same as the own elasticity of substitution with respect to health.

separable in preference. Unlike the preference formulations in the previous section, consumption and health are multiplicatively separable. It is assumed here that the degree of risk averse is the same across the population for theoretical convenience.

Since leisure is additively separable from consumption and health, the elasticities of substitution between consumption and leisure, and health and leisure are both equal to zero. It is easy to infer that the equilibrium condition for leisure retains similar properties as in the previous section where CARA preferences are assumed. Individual demand for leisure can be fully expressed in terms of the weighted village average leisure and a set of individual factors. When the population is homogenous, all individuals receive the village average leisure in equilibrium regardless of the severity of the health risks. What is not as transparent as the determination of leisure is the demand for consumption and sickness leave in equilibrium, because the elasticity of substitution between these two arguments are not equal to zero.

Let us maintain the assumptions on health risks, health status, time allocation, welfare weights, etc. as in the previous section, so that the results from these two types of preference formulations can be easily compared.

The program determining village-level full insurance is given by:

$$(3.22) \quad \underset{\{C_{is}, T_{is}, L_{is}\}}{Max} \quad \sum_{i=1}^{N} \lambda_i \sum_{s=1}^{S} p_s [U(C_{is}, H_{is}) + V(L_{is})] \ ,$$

subject to resource constraints and nonnegativity constraint as in the previous section.

Equilibrium Conditions

The first-order conditions for consumption, sickness leave, and leisure are as follows:

$$\text{(3.23)}\quad \lambda_i\gamma\left(C_{is}^{\gamma}H_{is}^{1-\gamma}\right)^{-\alpha}\left(C_{is}^{\gamma-1}H_{is}^{1-\gamma}\right)=\mu_s,$$

$$\text{(3.24)}\quad \lambda_i\sigma_i(1-\gamma)\left(C_{is}^{\gamma}H_{is}^{1-\gamma}\right)^{-\alpha}\left(C_{is}^{\gamma}H_{is}^{-\gamma}\right)=w_{is}\mu_s, \text{ and}$$

$$\text{(3.25)}\quad \lambda_i L_{is}^{-\beta}=w_{is}\mu_s.$$

Following the notations in Section 3.2.1, let us define $C_s=\sum_{i=1}^{N}C_{is}$, the total village consumption in state s; $T_s=\sum_{i=1}^{N}T_{is}$, the total village sickness leave in state s; $L_s=\sum_{i=1}^{N}L_{is}$, the total village leisure in state s. Rearranging the terms, we can rewrite individual consumption, sickness leave, and leisure in terms of the village aggregates, wage rate, health risks, welfare weights, and the parameters in preference as follows:

$$\text{(3.26)}\qquad C_{is}=\frac{\lambda_i^{\alpha-1}\left(\dfrac{w_{is}}{\sigma_i}\right)^{1-\gamma}}{\sum_{j=1}^{N}\lambda_j^{\alpha-1}\left(\dfrac{w_{js}}{\sigma_j}\right)^{1-\gamma}}\times C_s \text{ for all } i, j, \text{ and } s,$$

$$(3.27)\qquad L_{is} = \frac{\left(\frac{w_{is}}{\lambda_i}\right)^{\frac{1}{-\beta}}}{\sum_{j=1}^{N}\left(\frac{w_{js}}{\lambda_j}\right)^{\frac{1}{-\beta}}} \times L_s \text{ for all } i, j, \text{ and } s, \text{ and}$$

$$(3.28)\quad T_{is} = \frac{\frac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}{\sum_{j=1}^{N}\frac{\lambda_j^{\frac{1}{\alpha}} w_{js}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_j}} \times \left(T_s - \sum_{j=1}^{N}\varepsilon_{js}\right) + \left(\frac{\frac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}{\sum_{j=1}^{N}\frac{\lambda_j^{\frac{1}{\alpha}} w_{js}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_j}} \times \sum_{j=1}^{N}\frac{\overline{H}_j}{\sigma_j} - \frac{\overline{H}_i}{\sigma_i}\right) + \varepsilon_{is}$$

for all i, j, and s.

Equation (3.26) suggests that individual consumption is proportional to the village aggregate, and the proportion depends on the individual's welfare-weight adjusted wage rate relative to the sum of welfare-weight adjusted wage rates. In particular, an individual with a higher welfare weight or a higher wage rate will receive more consumption in equilibrium, controlling for other factors. Assuming equal welfare weights and wage rates, equation (3.26) is reduced to $C_{is} = \frac{C_s}{N}$, i.e., when the population is homogenous, consumption is equalized across population, which is usually referred to as "consumption smoothing" in a homogenous population.

The difference between (3.26) and (3.19) is that wage rate also enters in (3.26). This is because consumption and health are not additively separable as assumed in Section 3.2.2.

The term $\sum_{j=1}^{N} \lambda_j^{\alpha-1} \left(\frac{w_{js}}{\sigma_j} \right)^{1-\gamma}$ can be viewed as the weighted village average wage rates adjusted by risk attitudes and recuperation rate. An individual with a higher wage rate than that of this weighted average receives higher consumption in equilibrium, controlling for other factors.

Equation (3.28) describes the relationship between individual sickness leave and the village average sickness leave. The term $\left(T_s - \sum_{j=1}^{N} \varepsilon_{js} \right)$ has the interpretation of village aggregate net investment in health in state s. The common term $\dfrac{\dfrac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}{\sum_{i=1}^{N} \dfrac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}$ appears in all three terms in (3.28). It can be interpreted as a resource allocation factor in equilibrium. An individual whose welfare weight is higher than the weighted village average welfare weight receives longer sickness leave in equilibrium, controlling for other factors. An individual whose recuperation ability is higher than the weighted village average takes shorter sickness leave in equilibrium, controlling for other factors.

Wage rate plays a slightly different role in (3.28) than in (3.20). As in (3.20), (3.28) also implies that controlling for other factors, an individual whose wage rate is higher than the weighted village average wage rate receives shorter sickness leave in equilibrium because

$w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}$ is decreasing in wage rate (the exponent is negative). According to (3.20), an individual's share of the village aggregate net investment in health does not depend on his or her wage rate because individual wage rate enters in (3.20) additively. In (3.28), an individual's wage rate determines the magnitude of the resource allocation factor $\frac{\frac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}{\sum_{i=1}^{N} \frac{\lambda_i^{\frac{1}{\alpha}} w_{is}^{\frac{\gamma\alpha-\gamma-1}{\alpha}}}{\sigma_i}}$ along with other individual factors. An individual's share of the aggregate village net investment in health is negatively related with his or her wage rate. The fundamental difference in theoretical implications between (3.20) and (3.28) arise from the assumptions on preferences. From a reduced form perspective, (3.20) is derived from an additively separable preference specification, therefore wage rate enters in the determination of individual sickness leave in an additive fashion. On the other hand, (3.28) is derived from a multiplicatively separable preference specification, therefore wage rate affects an individual's sickness leave proportionally.

Conditional on welfare weights, recuperation rate, wage rate, and health risks, we learn from the second additive term in (3.28) that an individual with a higher natural health status will take shorter sickness leave in equilibrium. The last additive term in (3.28) implies that conditional on other factors, an individual afflicted with more severe health problem will take longer sickness leave in equilibrium. All of these suggest that health status will also be smoothed across possible states in equilibrium. In particular, when the population is large and homogenous, (3.28) is reduced to $T_{is} = \frac{T_s}{N} - \bar{\varepsilon} + \varepsilon_{is}$, i.e., the length of sickness leave varies with the size of the health risks and nothing else (average sickness leave $\frac{T_s}{N}$ and average sickness severity $\bar{\varepsilon}$ are known). In this case, individual's health status will be $H_{is} = \overline{H}_s + \sigma\left(\frac{T_s}{N} - \bar{\varepsilon}\right)$, entirely determined regardless of the severity of idiosyncratic health risks.

In summary, the theoretical implications under the multiplicative separable CRRA preference specification introduced in this section are very similar to those suggested by the additively separable CARA preference specification. The similarities and differences between these two specifications are illustrated in Table 3.1. Empirical strategies in testing for full insurance are discussed at the end of this chapter.

3.2.4 NONSEPARABILITY IN CONSUMPTION AND HEALTH

Additive separability and multiplicative separability provide similar theoretical predictions because they implicitly impose very strong restrictions on the substitution patterns across utility arguments. They have strong theoretical and empirical appeals because the equilibrium conditions derived are well-behaved and tractable. For example, by assuming separability in consumption and health, one implicitly imposes the assumption that the substitutability between consumption and health follows a fixed pattern for all levels of consumption and health. In the case of additively separable utility specification, marginal utility of consumption can increase without triggering any tradeoffs in health. As a result, the marginal rate of substitution between consumption and health in equilibrium always responds to the same magnitude with regard to changes in the relative levels of consumption and health.

A serious problem of the separability assumption on consumption and health is that it ignore the usual linkage of nutrition between consumption and health, which is not linear at all levels of consumption and health. Many empirical works have found that individual food consumption has positive impacts on health and labor productivity, but the impacts become very little after the individual's income level cross a benchmark level (see Ray, 1998, Dasgupta and Ray, 1990, Strauss and Thomas, 1998).

In the rural areas in developing countries, the problem of malnutrition is particularly severe. To fully understand the impacts of health risks on individual wellbeing, the nutrition link between consumption and health should be incorporated in preference specifications. This section introduces a preference specification in which health and consumption are nonseparable, i.e., the elasticity of substitution between consumption and health varies with the relative levels of consumption and health. Leisure is modeled as additively separable from consumption and health, since it is not of the key focus of this dissertation. Once nonseparability is involved, it is impossible to obtain well-behaved separate equilibrium conditions for individual consumption and health in terms of the corresponding village average and individual factors. We need to rely on certain linearization techniques to approximate the local properties of consumption and health around the optimum.

Let us start with the program determining Pareto optimality directly. The village-level full risk sharing is given by

$$(3.29) \quad \underset{\{C_{is},T_{is},L_{is}\}}{Max} \quad \sum_{i=1}^{N} \lambda_i \sum_{s=1}^{S} p_s [U(C_{is},H_{is})+V(L_{is})] \ ,$$

subject to the resource constraints and nonnegativity constraints (3.12)-(3.14). In equilibrium, the marginal rates of substitution across all utility arguments are equalized:

$$(3.30) \quad \frac{\partial U}{\partial C_{is}} = \frac{\sigma_i \partial U}{w_{is} \partial H_{is}} = \frac{\partial V}{w_i \partial L_{is}} = \mu_s .$$

Given the exact functional form of $V(\cdot)$, one can solve for the demand of individual leisure in terms of village average leisure and a set of individual-specific factors as in the previous two sections. Since the calculations are the same, they are not repeated in this section. Similar as leisure, the consumption-health composite is also smoothed across all states of nature, and can be written as the corresponding village average and a set of individual specific factors.

To understand how consumption and sickness leave are determined in equilibrium, let us focus on the first equality in (3.30):

$$(3.30)' \qquad U_C(C_{is},H_{is}) = \frac{\sigma_i}{w_{is}} U_H(C_{is},H_{is}) = \mu_s .$$

Denote the elasticity of the marginal utility of X with respect to Y as ξ_{XY} [3], and the percentage change in variable X as ΔX [4]. Log linearization of (3.30)' yields

$$(3.31) \qquad \xi_{HC}\Delta C_{is} \equiv \xi_{CH}\varsigma_T \times \Delta T_{is} + \Delta w_{is} + \xi_{CH}\varsigma_\varepsilon \times \Delta\varepsilon_{is} ,$$

where ς_T is the elasticity of health with respect to sickness leave, positive according to (3.1), and ς_ε is the elasticity of health with respect to health risks, a negative according to (3.1). Notice that all elasticity terms and percentage change terms are functions of the underlying utility arguments. The derivations are provided in Technical Appendix A.

---

[3] The elasticity of marginal utility of X with respect to Y is given by $\xi_{XY} = \dfrac{dU_X / U_X}{dY/Y} = \dfrac{YU_{XY}}{U_X}$.

The merit of (3.31) is that we are able to tell the local properties of consumption and sickness leave around optimum without having to know the actual levels of consumption and sickness leave. The sign of the elasticity of marginal utilities ξ_{CH} and ξ_{HC} depends on the sign of the cross partial derivative in $U(\cdot)$. When U_{CH} (or U_{HC}) is positive, the elasticity of marginal utilities are also positive. By $U_{CH} > 0$, we mean that good health improves an individual's marginal enjoyment from consumption, and higher consumption also translates into higher marginal enjoyment from health for the individual. The latter is the nutrition link between consumption and health mentioned earlier in this section.

Conditioning on wage rate and sickness leave, (3.31) suggests that when afflicted with illnesses or injuries, an individual's consumption decreases at rate $\frac{\xi_{CH}}{\xi_{HC}}\varsigma_{\varepsilon}$, which itself is also a function of consumption, health, and health risks:

$$(3.32) \quad \frac{\xi_{CH}}{\xi_{HC}}\varsigma_{\varepsilon} = -\frac{\sigma\varepsilon U_H}{CU_C}.$$

Conditional on other factors, (3.32) suggests that the rate that consumption drops due to health problems increases with the severity of the health problems $\sigma\varepsilon$. In (3.1) σ has the interpretation of recuperation rate when associated with sickness leave. When associated with health risks, it is interpreted as the rate of depreciation in health. The rate of consumption drop becomes lower if an individual starts with a higher consumption

[4] The percentage change of X_s is given by $\Delta X_s = \frac{X_{s'} - X_s}{X_s}$

level, as if consumption is used as a cushion to ward off health risks. In other words, there are tradeoffs between consumption and health in this case. It is important to realize that this cushioning or tradeoff only takes place in the case of nonseparability between consumption and health.

Wage rate plays a similar role in (3.31) as in (3.26) where preferences are assumed to be multiplicatively separable in consumption and health. Conditioning on health risks and sickness leave, an individual's consumption increases with wages at the rate $1/\xi_{HC}$, i.e., one percent of increase in wage rate leads to a ξ_{HC} percent of increase in consumption.

The percentage change of sickness leave and the percentage change of consumption are related by the ratio $\xi_{CH}\varsigma_T/\xi_{HC}$. Conditioning on wage rate and health risks, one percent increase in sickness leave leads to a $\xi_{CH}\varsigma_T/\xi_{HC}$ percent of increase in consumption. Similar to (3.32), this ratio can be written as

$$(3.33) \qquad \frac{\xi_{CH}}{\xi_{HC}}\varsigma_T = \frac{\sigma T U_H}{C U_C}.$$

Equation (3.33) describes how improved health status leads to higher demand for consumption. The rate of this health-driven demand for consumption increases with σT, the effectiveness of sickness leave, and decreases with one's own consumption status because of the law of diminishing marginal utility.

In (3.31), take the linear approximations of an individual's consumption around the village average consumption, we can relate an individual's consumption with the vilalge average consumption as in the previous sections:

$$(3.34)\quad \frac{C_i - \overline{C}}{\overline{C}} \cong \frac{\mu - \overline{\mu}}{\xi^i_{HC} \times \overline{\mu}} \text{ (see Technical Appendix B),}$$

where $\overline{\mu}$ represents the shadow price on the village budget constraint when the consumption of everyone is equal to the village average consumption. $\overline{\mu}$ is a function of the village aggregate economic resources and welfare weights.

Similarly, an individual's sickness leave can be expressed in terms of village average sickness leave, wage rates, and a set of individual-specific factors:

$$(3.35)\quad \frac{T_i - \overline{T}}{\overline{T}} \cong \frac{\mu - \overline{\mu}}{\sigma_i \xi^i_{CH} \times \overline{\mu}} + \frac{w_i - \overline{w}}{\sigma_i \xi^i_{CH} \times \overline{w}} + \frac{\varepsilon_i - \overline{\varepsilon}}{\overline{\varepsilon}} \text{ (see Technical Appendix B),}$$

where $\overline{w}$ denotes the average wage rate of the village, and $\overline{\varepsilon}$ denotes the average severity of health risks of the village. Notice that $\overline{\varepsilon}$ converges to the mean of the idiosyncratic health risks as the number of individuals goes to infinite.

(3.34) and (3.35) determine the following implicit functions of individual consumption and sickness leave:

$$(3.36)\qquad C_i = f^1\left(\overline{C}, \theta_i\right), \text{ and}$$

$$(3.37)\qquad T_i = f^2\left(\overline{T}, w_i, \varepsilon_i, \theta_i\right),$$

where θ_i represents a vector of individual characteristics such as risk attitude, welfare weights, and recuperation rate. Factors in θ_i do not have to be the same in both equations.

According to (3.34) and (3.35), an individual's consumption changes proportionally with changes in the average consumption of the village. Different individuals have different degree of comovement with the village average consumption. The proportion of comovement depends on ξ_{HC}, the substitutability between consumption and health, and the risk attitude of the individual. If there is a uniform risk attitude for all individuals, the comovement between an individual's consumption and the village average consumption is equal to 1. This is the same implication derived from the separable preferences. When ξ_{HC} for an individual increases, that is, when the individual becomes less and less risk averse, the degree of comovement between the consumption of this individual and the village average decreases, other things equal.

Similarly, an individual's sickness leave moves proportionally with the average village sickness leave. When all individuals have the same risk attitude, there is perfect comovement between the sickness leave of any individual and the village average sickness leave, other things equal. Controlling for other factors, an individual afflicted with more severe illness takes longer sickness leave. Due to the linear formulation of health determination in (3.1), the relationship between sickness leave and health risks are the same as in the separable preferences.

Other things equal, higher wage rate leads to longer sickness leave, namely, income effects dominate substitution effects, although different individuals respond differently. This is also a similar result from those derived under separable preferences.

3.3 CONCLUSION

When consumption and health are separable in preferences, they interact only through the village budget constraint by affecting the village aggregate labor income in equilibrium. Conditioning for the village aggregate labor income, health risks have no impact on an individual's consumption, and there is no tradeoff between consumption and health under full insurance. If insurance is imperfect, both consumption and health will be affected, and there will be tradeoffs between consumption and health. This suggests that when testing full insurance under separable preferences, we should estimate an individual's consumption and sickness leave jointly in a system such as

$$(3.38) \qquad \begin{cases} C_{it} = \alpha_i^1 + \beta^1 \overline{C}_t + \gamma_i^1 T_{it} + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} = \alpha_i^2 + \beta^2 \overline{T}_t + \gamma_i^2 C_{it} + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + v_{it} \end{cases}.$$

X_{it} and Y_{it} include variables such as household demographic characteristics and family wealth. u_{it} and v_{it} are error terms representing idiosyncratic health shocks. Testing full insurance in this case is to test whether or not the coefficients on X_{it} and Y_{it} equal to zero, because they are the factors that determine the scope and coverage of insurance an individual receives when the village-level mutual insurance is imperfect. The tradeoff between consumption and health due to imperfect insurance can be inferred from the coefficient on sickness leave in the consumption equation.

When consumption and health are nonseparable in preferences, health problems affect both consumption and health (sickness leave) even if an individual is under perfect insurance. Due to nonseparability, the consumption-health composite can be broken into two sequences, a consumption sequence that includes sickness leave on the right-hand side, and a sickness leave sequence that includes consumption on the right-hand side. The two sequences shall be estimated jointly similar as in (3.38). However, the inferences drawn here will be very different from the inferences drawn under the separable specifications.

Under full insurance, the coefficient on sickness leave in the consumption sequence and the coefficient on consumption in the sickness sequence describe the substitutability between consumption and heath due to nonseparability. Under imperfect insurance, however, these coefficients also reflect the tradeoff between consumption and health. If we reject full insurance under nonseparability, the tradeoff between consumption and health and associate welfare losses cannot be directly inferred from these coefficient. The following proposes a simple approach to this problem.

Let us consider a predicted value of consumption, $\hat{C}_{it}$, from a first-stage regression on a set of variables that are independent of health risks. The variations in $\hat{C}_{it}$ can be used to account for the part of consumption variations due to nonseparability between consumption and health. The variations in the remaining part of consumption,

$\left(C_{it} - \hat{C}_{it}\right)$, are therefore due to imperfect insurance. The coefficient on this residual term in the sickness leave sequence can be used to infer the tradeoff between consumption and health and the associated welfare losses in imperfect insurance. Similarly, a residual term $\left(T_{it} - \hat{T}_{it}\right)$ can be included in the consumption sequence. In other words, full insurance can be tested by the following:

$$(3.39)\qquad \begin{cases} C_{it} = \alpha_i^1 + \beta^1 \overline{C}_t + \gamma_i^1\left(T_{it} - \hat{T}_{it}\right) + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} = \alpha_i^2 + \beta^2 \overline{T}_t + \gamma_i^2\left(C_{it} - \hat{C}_{it}\right) + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + v_{it} \end{cases}$$

The coefficients on X_{it} and Y_{it} can be used to test full insurance as in (3.38). The coefficients on the residual terms can be used to test full insurance, as well as to infer the tradeoff between consumption and health and welfare losses due to imperfect insurance.

Table 3.1 summarizes the theoretical predictions from the three preference specifications. Take, for example, two individuals, *A* and *B*, residing in the same village, encountering health shocks ε_A and ε_B respectively. Let us assume that they have the same initial health status prior to the health shocks. *A* and *B* will take sickness leave T_A and T_B so that they are able to recover to H_A and H_B respectively. Under full insurance, other things equal, they both can recover to the same level of health, regardless of the magnitude of the health shocks. That is, both will take necessary and sufficient measure to recover from illnesses, as if there exists an implicit reimbursement system in the village to pay for an individual's health care and cover all income losses associated with seeking health care. In other words, health is smoothed across all possible states for *A* and *B*. One way to test full insurance is to compare the health outcomes of *A* and *B* to see whether they are the same controlling for other factors, if health outcomes are available in the data.

Similar arguments can be made about the other socioeconomic factors appear in the models. For example, under separable preferences, the consumption for both individuals reach the same level regardless of the amount of sickness leave they take, the severity of health risks they encounter, and their initial health status, as if there exists an implicit income redistribution channel through which the consumption for each individual are equalized across all states of nature and the consumption for both individuals are equalized for any given state. Under nonseparable preferences, the consumption for both individuals are affected by the amount of sickness leave they take, but this is compatible with full insurance because the substitutability between consumption and health varies with the levels of consumption and health for each individual. For example, if individual *A* has a higher initial health status or higher wage rate than individual *B*, *A* will receive higher consumption and lower sickness leave in equilibrium than *B*.

Figure 3.1 General Utility Tree

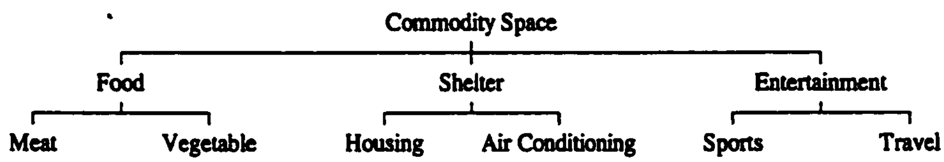

Figure 3.2 Separability in Consumption and Health

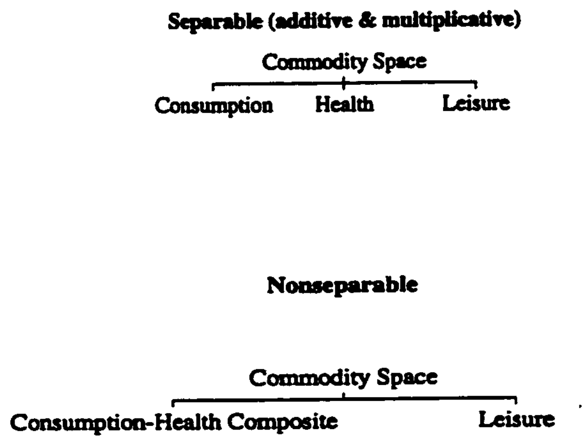

Table 3.1 Three Preference Specifications and the Theoretical Predictions

	Consumption			Sickness Leave		
Variable	Additively Separable	Multiplicatively Separable	Non-separable	Additively Separable	Multiplicatively Separable	Non-separable
Aggregate Consumption	+	+	+			+
Aggregate Sickness Leave			+	+	+	+
Individual Sickness Leave			+			+
Individual Consumption						+
Health Risks			–	+	+	+
Recuperation Rate			+	–	–	–
Natural Health Status		+	+	–	–	–
Risk Aversion	+	+	–	+	+	–
Wage Rate		+	+	–	–	–
Welfare Weight	+	+	+	+	+	+

Note: "+" represents an increasing relationship conditioning on other factors. "–" represents a decreasing relationship, conditioning on other factors. The shaded area means that there is no interaction between variables as predicted by the theory, other things equal.

TECHNICAL APPENDIX A

The theorems presented in the following are from Luenberger (1995).

A.1. INDEPENDENCE OF PREFERENCES

Independence theorem. Suppose a preference ordering is represented by a utility function $u(y,z)$. Then if y is independent of z, u can be written as $u(y,z) = U(v(y),z)$ where $U(v,z)$ is strictly increasing in v. Furthermore, if u is continuous and strongly monotonic, then v and U are continuous and v is strongly monotonic. Conversely, if a preference ordering is represented in the form able, they y is independent of z.

This theorem is very useful since it shows that if y is independent of z there is a special utility function for the y commodities. Its value, together with z, is all that matters in the determination of the overall utility.

Independence of preferences. Suppose there are a total of m commodities partitioned into P groups. A partition is defined by $\{N_1, N_2, \ldots, N_P\}$, with $N_i \cap N_j$ empty for $i \neq j$, and $\cup_{i=1}^{P} N_i = N$. An arbitrary commodity bundle $x \in R_+^m$ is partitioned correspondingly as $\mathbf{x}=(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_P)$. Let $\mathbf{x}_{-i}$ denote the vector of commodities in the complement to N_i. That is, $\mathbf{x}_{-i}=(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_{i-1},\mathbf{x}_{i+1},...,\mathbf{x}_P)$. A preference ordering is

(a) Weakly independent with respect to the partition $\{N_1,N_2,...,N_P\}$ if for every i =1, 2, …, P the vector $\mathbf{x}_i$ is independent of its complement $\mathbf{x}_{-i}$;

(b) Strongly independent with respect to the partition $\{N_1,N_2,...,N_P\}$ if it is weakly independent with respect to $\{N_1,N_2,...,N_P\}$, and with respect to the partitions consisting of all unions of $N_1,N_2,...,N_P$ that are proper subsets of N.

Stronger structural forms are obtained by introducing additional independence relations. The most important are the weak and strong versions of independence relative to a partition. Separability of utility.

A.2 SEPARABILITY OF UTILITY

Parallel with the above definitions for preference orderings, there are similar definitions for utility functions.

Separability of utility. A continuous utility function u is

(a) Weakly separable with respect to the partition $\{N_1, N_2, ..., N_P\}$ if there are continuous functions U and $u_1, u_2, ..., u_P$ such that for all $x \in X$

$$u(x) = U(u_1(x_1), u_2(x_2), ..., u_P(x_P));$$

(b) Strong separable with respect to $\{N_1, N_2, ..., N_P\}$ if there exist continuous functions U and $u_1, u_2, ..., u_P$ such that for all $x \in X$

$$u(x) = U\left(\sum_{i=1}^{P} u_i(x_i)\right)$$

Proposition. Suppose a preference ordering is continuous and strongly monotonic. Suppose the variables are partitions into P groups. Then the preference ordering is weakly independent with respect to this partition if and only if every continuous utility function that represents it is weakly separable. (For proofs, see Luenberger, 1995)

This result has great intuitive appeal. It says that under weak independence each commodity group can be assigned its own utility function. Within each group, there may also exist weakly separable subgroups. A budget can be allocated to a group and then further be allocated to the subgroups. In our case, if consumption, health, and leisure are weakly separable in preference, an individual's income is allocated first to these three groups, then for each group, say consumption items, the income can be further allocated separately to food, clothes, traveling, etc., if these are weakly separable within consumption as well.

Strong independence theorem. Suppose a preference ordering is continuous and strongly monotonic. Suppose the variables are partitioned into P groups with $P > 2$. Then the preference ordering is strongly independent with respect to this partition if and only if every continuous utility function which represents it is also strongly separable.

A very special case of strong separability of utility functions occurs when the partition consists of the element-by-element partition. In this case the utility function is equivalent to an additive utility function.

$$u(x)=\sum_{i=1}^{P} u_i(x_i).$$

TECHNICAL APPENDIX B

B.1 ELASTICITY

In a general health investment function

$$\text{(B.1)} \quad H = f(\overline{H}, T, \varepsilon),$$

the elasticity of health with respect to sickness leave is given by

$$\text{(B.2)} \quad \varsigma_T = \frac{dH/H}{dT/T} = \frac{TH_T}{H},$$

and the elasticity of health with respect to illness is given by

$$\text{(B.3)} \quad \varsigma_\varepsilon = \frac{dH/H}{d\varepsilon/\varepsilon} = \frac{\varepsilon H_\varepsilon}{H}.$$

In the case of (3.1), the net investment in health is linear,

$$\text{(B.4)} \quad f(\overline{H}, T, \varepsilon) = \overline{H} + \sigma(T - \varepsilon),$$

we have $\varsigma_T = \frac{\sigma T}{H}$, and $\varsigma_\varepsilon = \frac{\sigma\varepsilon}{H}$.

In a general utility function $U(C, H)$, the elasticity of marginal utility of consumption with respect to health is given by

$$\text{(B.5)} \quad \xi_{CH} = \frac{dU_C/U_C}{dH/H} = \frac{HU_{CH}}{U_C},$$

the elasticity of marginal utility of health with respect to consumption is given by

$$(B.6) \quad \xi_{HC} = \frac{dU_H / U_H}{dC / C} = \frac{CU_{CH}}{U_H},$$

and the elasticity of substitution between consumption and health is given by

$$(B.7) \quad \zeta_{CH} = -\frac{d(U_C / U_H)/(U_C / U_H)}{d(H/C)/(H/C)}.$$

B.2 LOG LINEARIZATION AND THE FIRST-ORDER CONDITIONS:

The first-order Taylor expansion of function $f(X)$ around X_0 is given by

$$(B.8) \quad f(X) \cong f(X_0) + (X - X_0) \times f_X \mid_{X_0}.$$

If $f(X) = \log X$, we have

$$(B.9) \quad \log X \cong \log X_0 + \frac{(X - X_0)}{X_0}.$$

Define $\Delta X_0 = \frac{(X - X_0)}{X_0}$, (B.9) can be written as

$$(B.10) \quad \log \frac{X}{X_0} \cong \Delta X_0,$$

that is, the small percentage change in variable X (the growth rate of X in a continuous case) can be approximated by the logarithm of the relative ratio between two levels.

In Chapter 3, we know that the marginal utility of consumption is equalized to the marginal utility of health adjusted by wage rate and recuperation rate. That is, $U_C = \frac{\sigma}{w} U_H$ for all $\{C, H, w\}$. Therefore,

$$U_C = \frac{\sigma}{w} U_H = \mu \tag{B.11}$$

$$U_{C'} = \frac{\sigma}{w'} U_{H'} = \mu' \tag{B.12}$$

Recall that μ is the state-dependent shadow price of the village aggregate budget constraint.

Divide (B.11) by (B.12), we have

$$\frac{U_{C'}}{U_C} = \frac{w}{w'} \frac{U_{H'}}{U_H} = \frac{\mu'}{\mu}. \tag{B.13}$$

Take logarithm on both sides of first equality in (B.13), approximate around $\{C, H, w\}$, using (B.10), we have

$$\Delta U_C \cong \Delta U_H - \Delta w. \tag{B.14}$$

From (B.5), (B.6), and (B.10), (B.14) can be written as

$$\xi_{CH} \Delta H \cong \xi_{HC} \Delta C - \Delta w. \tag{B.15}$$

Given (B.2)-(B.4), $\Delta H \cong \varsigma_T \times \Delta T + \varsigma_\varepsilon \times \Delta \varepsilon$, (B.15) can be rewritten as

$$(3.27) \qquad \xi_{HC}\Delta C_{is} \cong \xi_{CH}\varsigma_T \times \Delta T_{is} + \Delta w_{is} + \xi_{CH}\varsigma_\varepsilon \times \Delta \varepsilon_{is}.$$

B.3 INDIVIDUAL CHOICE VARIABLES AND VILLAGE AGGREGATE VARIABLES UNDER NONSEPARABILITY

Generally, an individual's choice variables can be related to the corresponding village aggregate choice variables and a set of parameters representing individual-specific factors. Such calculations are usually carried out in three steps. Take, for example, individual consumption. In the first step, an individual's consumption is expressed in terms of individual factors and the shadow price associated with the village aggregate budget constraint. In the second step, sum up consumption for all individuals to construct the village aggregate consumption, which is expressed in terms of individual-specific factors across the village and the shadow price. Finally, replace the shadow price in the first expression by village aggregate consumption and individual-specific factors for all individuals based on the results from the second step. At this point, an individual's consumption is related to his or her characteristics, the characteristics of others within the village community, and the village aggregate consumption.

In a similar fashion, (3.27) can be used to derive the relationship between individual choice variables and the corresponding village aggregate variables under a general nonseparable preference.

According to (B.10), the first-order condition (B.13) can be approximated around $\{C, H, w\}$ through log linearization as

$$\text{(B.16)} \qquad \Delta U_C \cong \Delta U_H - \Delta w = \Delta\mu .$$

Using (B.2)-(B.4), (B.16) can be written in terms of elasticities and percentage changes in consumption and sickness leave as

$$\text{(B.17)} \quad \xi_{HC}\Delta C_{is} \cong \xi_{CH}\varsigma_T \times \Delta T_{is} + \Delta w_{is} + \xi_{CH}\varsigma_\varepsilon \times \Delta\varepsilon_{is} \cong \Delta\mu ,$$

which implies that

$$\text{(B.18)} \qquad \Delta C_i \cong \Delta\mu / \xi_{HC}^i \text{, and}$$

$$\text{(B.19)} \qquad \Delta T_i \cong \left(\Delta\mu - \Delta w_{is} - \xi_{CH}\varsigma_\varepsilon \times \Delta\varepsilon_{is}\right)/\xi_{CH}\varsigma_T .$$

Denote the village average consumption and sickness leave as $\overline{C}$ and $\overline{T}$. Denote the average village wage rate as $\overline{w}$. If we approximate individual i's consumption and sickness leave around the village average values, (A.18) becomes

$$\text{(B.18)'} \qquad \frac{C_i - \overline{C}}{\overline{C}} \cong \frac{\mu - \overline{\mu}}{\xi_{HC}^i \times \overline{\mu}},$$

where $\overline{\mu}$ represents the shadow price on the village budget constraint when the consumption of everyone is equal to the village average consumption.

Similarly, (B.19) becomes

$$\text{(B.19)'} \qquad \frac{T_i - \overline{T}}{\overline{T}} \cong \frac{\mu - \overline{\mu}}{\sigma_i \xi_{CH}^i \times \overline{\mu}} + \frac{w_i - \overline{w}}{\sigma_i \xi_{CH}^i \times \overline{w}} + \frac{\varepsilon_i - \overline{\varepsilon}}{\overline{\varepsilon}},$$

where $\bar{\varepsilon}$ represents the average severity of health risks in the village. $\bar{\varepsilon}$ converges to the true mean of the health risks as the number of individuals in the village goes to infinite.

(B.18)' and (B.19)' can be put into reduced-form estimations similar to (3.33) as follows

$$\text{(B.20)} \quad \begin{cases} C_{it} - \overline{C}_t = \alpha_i^1 + \beta^1 \overline{C}_t + \gamma_i^1 T_{it} + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} - \overline{T}_t = \alpha_i^2 + \beta^2 \overline{T}_t + \gamma_i^2 C_{it} + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + \eta_i (w_{it} - \overline{w}_t) + v_{it} \end{cases}$$

Rearranging the terms, (3.33) can be turned into

$$\text{(3.34)} \quad \begin{cases} C_{it} = \alpha_i^1 + \beta^1 \overline{C}_t + \gamma_i^1 (T_{it} - \hat{T}_{it}) + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} = \alpha_i^2 + \beta^2 \overline{T}_t + \gamma_i^2 (C_{it} - \hat{C}_{it}) + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + v_{it} \end{cases}.$$

CHAPTER 4 SELECTION PROBLEMS IN THE ICRISAT LABOR DATA

Missing observations is a typical problem associated with cross-sectional time-series data. If observations are missing due to purely random reasons and if the sample size is sufficiently large, using only the complete observations in estimations will not affect the consistency of predictions, although standard errors may be inconsistent if not properly treated. However, in most cases, observations are missing either because of censoring or selection, and using only the complete observations will no longer be consistent (Greene 1994, Hsiao 1986). As many researchers have noticed, the ICRISAT labor data are noisy and sometimes even awkward to use (Townsend 1994, Kochar 1999). This essay focuses on the problem of missing observations in the ICRISAT labor data. Two classes of discrete choice models are applied to explore the possible reasons for missing observations, sample selection, and the potential biases in estimates if complete observations are used in estimations. The results are brought into empirical tests for full insurance in another essay of the dissertation, titled "Risks, Insurance, and the Decision to Take Sickness Leave: Townsend Revisited in General Equilibrium."

4.1. MISSING OBSERVATIONS IN THE ICRISAT LABOR DATA

The ICRISAT labor data was collected continuously throughout the survey years but coded quite differently for 1975-1977 and 1978-1984. The earlier labor data contains more details on labor hours by activity and source. It identifies production activities such as crop production, animal husbandry, building, repairs, trade, domestic work, food and fuel gathering and others. Data on total working hours and wage income were collected for each village labor market participant by these activities. The data also contain a series of binary indicators on the loss of work since the last interview due to reasons such as illnesses, injuries, involuntary unemployment, travel, religious events, etc. The later labor data only distinguishes activities by the nature of employment—farm work, off-farm non-government work, and off-farm government work. Information on employment days, average working hours per day and total wage income received in cash and kind for each type of employment is provided. Hours spent on food and fuel gathering were not recorded nor were hours spent on domestic work. Information on the loss of work is in the same format as before. Information on village labor market participants' characteristics such as age, sex, education, caste, and relation with household head is available from a census data collected in the early years of the survey.

The investigator interviewed labor market participants at three-to-four-week intervals to collect information on labor activities on a recall basis. Ideally, we should see a continuous panel of labor data 11 to 14 times a year for each interviewee. If one were to look at the actual frequencies the labor market participants were interviewed, the picture is mixed. Only around 70 percent of the interviewees had a regular continuous panel of 11 to 14 times a year. About 12 percent of the interviewees had shorter panels, between 1 to 10 times a year. The rest had longer panels, between 15 to 23 times a year. It is quite common that information on a same individual was collected at different frequencies across years, and that information on different individuals were collected at different frequencies in a same year. In a typical year, around 12.6 percent of the data have missing entries and no explanation was provided as for why. For example, information on an individual's labor activities could be completely missing at a certain point of time, and the binary indicators of the reasons for the loss of work also not available. Therefore, it would be impossible for us to identify whether missing entries resulted from human errors or from some important economic decisions that the individual had made.

There are exogenous as well as endogenous reasons as for why information on a labor market participant is missing at a particular point of time. It could result from errors in interview and data processing. For example, the investigator was not able to interview the individual in the first visit and failed to go back for a follow-up interview. It may also be a result of self-selection, that is, the investigator had made every effort, and it was the interviewee who had adjusted his or her socioeconomic status that led to missing observations in the data.

This problem has not been properly dealt with in empirical works using the ICRISAT labor data. For example, Townsend (1994) uses only 1978-1984 male labor data because the early labor survey was of a different format, and information on women and children "seem especially unreliable" (Townsend 1994, page 546). He obtained adult male wage series by dividing the recorded wage income by the recorded number of days worked and then averaging across interviewed adult male laborers. He then derived the village average labor supply by averaging the labor supply of all interviewed adult male laborers in a household and then averaging across households. The average village labor supply is used to derive average leisure and the proportion of year sick for each adult male laborer, which are key explanatory variables in testing full insurance. At no point of time is the problem of missing observations addressed in Townsend (1994). If observations are missing due to random reasons and the sample sufficiently large, the wage series and time allocation of the adult male laborers derived by using complete observations converge to the true values. If observations are missing due to self-selection, these measures will no longer be good approximations for the true values. For instance, when an adult male laborer dropped out of the village labor market temporarily to pursue other interests to support his family at the time of the interview, his information would be missing because he would not be considered as a participant on the village labor market. In the meantime, he might get sick and would choose to take sickness leave, but it would be impossible for us to know from the data. If this is the case, Townsend's imputation of the wage series of adult male laborers will have an upward bias, where as the imputation of average leisure and the proportion of year sick will have downward biases. When testing for full insurance where adult equivalent consumption is regressed on these imputed explanatory variables respectively in a univariate fashion, and Townsend found that "... household consumptions are not much influenced by ... sickness ... or other idiosyncratic shocks, controlling for village consumption" (Townsend 1994, page 539). If part of the missing entries were results of self-selection, Townsend's estimates on the proportion of sickness leave and average leisure will have a downward bias, whereas the estimates on wage rate will have an upward bias.

Another important work using the ICRISAT labor data is Kochar (1999). Kochar used the full ten-year labor data to examine intrahousehold allocations of labor hours across farm and off-farm activities and found evidence in support of labor diversification in income smoothing. She started with a two-stage dynamic programming problem in which farmers choose the level of labor input in the planting season (the first stage) without realizing any marketable output, and choose the level of labor input in the harvesting season (the second stage) while facing weather and other income shocks. Using the nonmissing observations on labor utilization, she found that laborers' working hours would increase in response to negative crop income shocks. The problem of missing observations was not addressed in her work. Following the previous arguments that missing observations might be a result of self-selection, it is possible that her findings have overemphasized the importance of labor diversification as a measure to cope with risks, because the estimates were based on those who are more active on the labor market.

We can see that without understanding the nature of the missing observations first, any further pursuit of empirical research using the ICRISAT labor data may lead to biases in estimates. In the following, two classes of discrete choice models are applied to explore the possible reasons of missing observations.

4.2. POISSON REGRESSIONS FOR INTERVIEW FREQUENCIES

The merit of using a repeated survey is its richness in time-series which allows us to follow up individual responses throughout the course of a time span, and its cross-sections which provides insights into group-wide heterogeneity. Missing observations make it difficult to keep track of individual responses across time and group, and interpolations are only applicable when missing observations are purely of random nature. To obtain a measure of how observable a labor market participant is, a series of Poisson regressions are applied in the following.

Poisson distribution is used to describe the incidence of a discrete event over a duration (or *exposure*). It is derived as a limiting case of the binomial distribution. A standard generalization of the Poisson is the negative binomial distribution. The Poisson regression model stipulates that each dependent variable y is drawn from a Poisson distribution with parameter λ, which is related to the regressors X. The most common formulation is

$$(4.1) \quad E[y \mid X] = \exp(X\beta).$$

If we want take a labor market participant's observed interview frequency as dependent variable and examine the underlying determinants of the observability, a Poisson regression would seem quite natural and appropriate because of its discrete and nonnegative nature.

Since the ICRISAT labor data is longitudinal, simple cross-sectional formulations will not be consistent because there may be unobserved individual-specific characteristics. A usual formulation in cases such as this would involve a parameter assumed to follow certain statistical distributions. For example, let y_{it} denotes individual i's counts of nonmissing entries in year t, we can assume that

$$(4.2) \quad y_{it} \sim Poisson(\alpha_i \lambda_{it}),$$

where $\lambda_{it} = \exp(X'_{it}\beta)$, for $i = 1,\dots,n$, $t = 1,\dots,T$, and α_i is the individual-specific unobserved characteristics. We may either treat α_i as a fixed effect constant across time, or a random effect follows a known distribution.

The choice between fixed effects and random effects is not an easy one. Generally, the random-effects model is appropriate if the sample is drawn from a single population and one wants to obtain inference on the population. It more easily accommodates random slope parameters as well as random intercepts. Consistency of the random-effects model is based on a rather strong assumption that the unobserved individual-specific characteristics are uncorrelated with the observed characteristics. The fixed-effects model makes no such assumption. It is appropriate if one only wants to explain the sample. All time-invariant regressors along with their coefficients are absorbed into the individual-specific effect and will not be identifiable from the constant. In this case, one cannot make any inferences on individual characteristics such as sex, education, and caste, which may of great theoretical and empirical interests. Hausman specification tests are available for random-effects and fixed-effects Poisson regressions.

In the following, both fixed- and random-effects specifications are employed. The results are presented in Table 4.2.a and Table 4.2.b respectively. A Hausman test is constructed based on the two estimators.

4.2.1 FIXED-EFFECTS POISSON

Let us start by assuming that when an investigator conducted interviews in a village to collect labor data, there were unobservable (to the econometrician) individual-specific time-invariant factors that affected the counts of the nonmissing data entries of an individual. In other words, we assume here that there are individual-specific fixed effects that affect an interviewee's observability in the labor data. Specifically, let α_i denote individual i's unobserved characteristics, and y_{it} individual i's counts of nonmissing data entries in year t.

If the sample size is small the model can be easily estimated by

$$(4.3) \quad \exp\left(\sum_i \alpha_i d_{it} + X'_{it}\beta\right),$$

where d_{it} is a dummy variable for individual i at date t. However, as the sample size increases, the number of regressors will increase, so that software restriction on the maximum number of regressors is likely to be encountered. Another potential problem is possible inconsistency if T is small and $n \to \infty$. These complications have made it unappealing to estimate the fixed-effects Poisson model directly. Instead, a conditional maximum likelihood approach is favored because it manages to cancel out the effects α_i

by formulating a conditional[1] joint-density (see Cameron and Trivedi 1998). The conditional log-likelihood function is given by

$$(4.4)\quad L=\sum_i\left[\log\left(\sum_t y_{it}\right)-\sum_t\log(y_{it}!)+\sum_t y_{it}\log\left(\frac{\exp(X_{it}'\beta)}{\sum_t\exp(X_{it}'\beta)}\right)\right].$$

This formulation is quite similar to that of the multinomial logit model, except that y_{it} is not restricted to taking only values zero or one and to sum over t to unity. The results from fixed effect Poisson are reported in Table 4.2.a. Notice that when fixed-effects is assumed, all individual-specific time-invariant characteristics are not identifiable from the constant. Coefficients on sex, occupation, caste, landholding class, etc. are suppressed.

We can see from Table 4.2.a that conditional on individual-specific fixed effects, interviewee's demographic factors such as age-sex group, education, and relation with household head have significant impacts on the individual's observability in the labor data. In particular, males of age 15 to 44 were more observable than females of similar age, so were the males of age 45 to 64, only to a lesser extent. This seems to be compatible with the general understandings of labor market participation in rural areas in developing countries. Males tend to dominate in most farm works due to the high physical demand. It is also compatible with the local religious and social customs that women are not allowed to participate in certain farm works such as plowing. The

[1] Conditional on $\sum_t y_{it}$.

difference in the estimates on the two age-sex groups show that women of child-bearing age participate even less than other adult women, controlling for other factors.

Estimates of individual and household education level suggest some evidence of self-selection in the survey. Controlling for other factors, interviewees with at least primary schooling were less frequently observed in the labor data possibly because they had better opportunities to support themselves as well as their families than working on the village labor market from time to time. On the other hand, individuals from a family with more educated members were more frequently interviewed. It is highly possible that educated family members were seeking outside opportunities, which in turn increased the family's need for labor income to offset risks associated with the outside opportunities. These can be viewed as diversification in household incomes in order to smooth consumption and welfare across time and states of the nature. They are economic decisions made by the households and individuals.

Another piece of key evidence supporting self-selection is that fitted per capita income obtained from a first-stage regression has a negative and significant impact on an interviewee's observability in the village labor market. Conditional on individual and household characteristics, when per capita income increases, household and individual's reliance upon incomes from the village labor market will decrease. As a result, individuals from higher income families participate less actively in the village labor market, either because they have better paid jobs to do, or because they have higher reservation wage than the their neighbors.

Year dummies are included in the regression to count for possible omitted time-varying variables such as rainfall, government-funded projects, and inflation variables, which affect villagers' participation in the village labor markets. They also serve as a check on whether there was some learning-by-doing of the investigator. 1975 is the omitted year. Compared to 1975, estimates of all of the following years are positive and significant. We know that annual rainfalls in the study villages were erratic, especially in Aurepalle and Shirapur. Government project openings were rare during the whole survey years. Inflation was mild in the early years of the survey, and was on a rising trend during 1978-1981, then stayed stagnant in the later years. If they had significant impacts on the villagers' labor participation, the estimates of the year dummies should reflect these rather nonsystematic trends, especially inflation since rainfall and government job openings were rather unexpected random events, and their effects on household decisions would be even out through the years. In other words, should the impacts from these external economic conditions dominate the investigator's learning-by-doing, we expect to see that the estimates largely follow the changes in price levels. However, the estimates of the year dummies seem rather steady, therefore, we infer that the investigator's learning-by-doing is a more plausible rationale. In fact, all of the ICRISAT investigators resided in the village throughout the survey years in order to establish good connections with the sample households, so that more subtle questions could be tackled, and information updated readily and easily. It is reasonable to believe that the investigators could have learned more about the households through the years, obtained more experience on handling data and dealing with missing observations.

4.2.2 RANDOM-EFFECTS POISSON

In this section, a random-effects Poisson model is employed to explore the possible reasons of missing observations. The results are presented in Table 4.2.b and comparable to those from the previous section.

Following the notations introduced in Section 2.1, let y_{it}^{h} denote the counts of individual i's nonmissing data entries in year t, and X_{it}^{h} denote his or her socioeconomic characteristics such as age, sex, relation to household head, fitted per capita income obtained from a first-stage regression, landholding class, etc. Let us assume that there exist a set of unobservable (to the econometrician) characteristics α_i that affect an interviewee's observability in the labor data, and the α_i's follow a gamma distribution[2] of parameter θ, $gamma(\theta,\theta)$, so that $E[\alpha_i]=1$ and $V[\alpha_i]=\frac{1}{\theta}$. Notice that the gamma distribution for integers is $\frac{\lambda^y}{y!}e^{\lambda\alpha}\alpha^{y-1}$, which is reduced to Poisson if $\alpha=1$. Therefore, it can be viewed as in the same distribution family as the Poisson, which gives us great computational ease because it allows us to derive a closed-form likelihood function.

---

[2] Alternatively, one may consider normally distributed random effects. However, the likelihood function does not have a closed form, approximations by a well-behaved polynomial will be needed in computations. In addition, the applicability of such approximations is usually restricted to small panel size.

The joint density function of individual i's annual counts of nonmissing data entries throughout the survey years is given by

$$(4.5)\ \Pr ob(y_{i1},\ldots,y_{iT}) = \left[\prod_t \frac{\lambda_{it}^{y_{it}}}{y_{it}!}\right] \times \left(\frac{\theta}{\sum_t \lambda_{it} + \theta}\right)^{\theta} \times \left(\sum_t \lambda_{it} + \theta\right)^{-\sum_t y_{it}} \frac{\Gamma\left(\sum_t y_{it} + \theta\right)}{\Gamma(\theta)}.$$

Imposing heterogeneity in the random effects, i.e., $\alpha_i \sim gamma(\theta^i, \theta^i)$, the log likelihood function for all individuals across time (See Greene 1997, Stata 6.0 Releases) can be written as

$$(4.6)\quad \begin{aligned} L = \sum_{i=1}^{N} \varpi_i \{ & \log \Gamma\left(\theta^i + \sum_{t=1}^{T} y_{it}\right) - \sum_{t=1}^{T} \log \Gamma(1 + y_{it}) - \log \Gamma(\theta^i) + \theta^i \log u_i \\ & + \log(1 - u_i) \sum_{t=1}^{T} y_{it} + \sum_{t=1}^{T} y_{it}(X_{it}\beta) - \left(\sum_{t=1}^{T} y_{it}\right) \log\left(\sum_{t=1}^{T} \lambda_{it}\right) \} \end{aligned},$$

where $u_i = \dfrac{\theta^i}{\theta^i + \sum_{t=1}^{T} \lambda_{it}}$, and $\lambda_{it} = \exp(X_{it}\beta)$.

If this is indeed the correct specification and if there is no human error in data processing, y_{it}^{h} has the interpretation of average labor market participation at the time of survey. As in the previous section, year dummies are included in the regression to serve as an account for possible omitted variables, as well as a check on whether the investigator had accumulated experience in data collection and processing.

Using the maximum number of a household's annual interviews as exposure, assuming individual-specific random effects, the regression results are provided in Table 4.2.b. Standard errors reported are robust to the grouping at household level. Unlike fixed-effects, individual-specific time-invariant factors are identifiable from the constant.

The results show that an individual's characteristics such as age, sex, education, and occupation have significant impacts on his or her observability in the labor data. Similar to the findings in fixed-effect Poisson regression, able-bodied males are more observable in the labor data than females of similar age, and the difference between females of child-bearing age and males in the same age group is larger than the 45-64 age-sex groups. Estimates of interviewee's own education level and the education level of household members are also similar to those from the fixed-effects Poisson regression. Individuals whose main occupation was agricultural are more observable in the labor data, because they participated more actively than those who pursued alternative occupations[3].

Relationship with the household head has a negative effect on an interviewee's observability. A direct member is defined as either the spouse or an offspring of the household head. About one-half of the sample households had extended family members, such as relatives, friends or permanent servants residing in the household. The relationship between an extended member and the rest of the household is presumably

[3] Village residents also participate in other income-earning activities such as trade, handicrafts, shopkeeping, moneylending, transport, buildings, and repairs, although the bulk of income was generated from crop production and the labor market.

loose and temporary in nature. Nonetheless, the extended families members are contributing income and sharing resources as well as risks with other members of the family. There was an increasing trend that families with extended structures would split into nuclear families over time (see Walker and Ryan 1990), which might have resulted in less frequent interviews with the extended family members.

Individuals from larger families were less frequently interviewed conditional on other factors, possibly due to the intrahousehold allocation of labor activites as discussed in the previous section. Similar to the results from fixed-effects Poisson regression, individuals from families with more educated members are more observable in the labor data. The interpretations are similar as in the previous section. Fitted per capital income is negative and significant as in the fixed-effects Poisson regression. These results all seem to be in favor of the self-selection hypothesis.

Caste is a very important factor in the social and economic lives in India, however, the estimate here is not significant, which is also compatible with findings in other researches (see Walker and Ryan 1990). It has been argued that most of labor contracts in the village are renewed on a daily basis, so that discrimination based on caste rankings cannot sustain. Caste rankings are also found to have very little impact on an employer's screening of employees, or an employee's choice of employers.

4.3 HAUSMAN SPECIFICATION TEST

As discussed earlier, the choice between random-effects and fixed-effects formulation is not an easy one. Both have their own strengths and weaknesses when dealing with unobserved characteristics. If we adopt the view that the unobserved effects are random from the outset and if the model is perfectly specified, the results from fixed-effects formulations can be viewed as conditional on the effects that are in the sample, and the results from random-effects formulations as unconditional inferences for the whole population (see Hsiao 1986). Even though fixed-effects and random-effects can be unified as such, when dealing with discrete data, inefficiencies and inconsistencies in estimations will still arise due to potential cross-sectional heterogeneities and serial correlations, therefore, checking for the correct specification is still of great importance.

Hausman specification test is based on the null hypothesis that random effects are uncorrelated with the regressors. If the null is true, both fixed- and random-effects estimators are consistent, but random-effects estimator is more efficient; otherwise the random effects estimator is inconsistent. Hausman tests for Poisson regressions use both estimators and their variance-covariance matrix to construct a chi-squared statistic (see Cameron and Trivedi 1998, Hsiao 1986):

$$(4.7) \quad T_H = \left(\hat{\beta}_{RE} - \tilde{\beta}_{FE}\right)' \left[V\left(\tilde{\beta}_{FE}\right) - V\left(\hat{\beta}_{RE}\right)\right]^{-1} \left(\hat{\beta}_{RE} - \tilde{\beta}_{FE}\right),$$

where $\hat{\beta}_{RE}$ denotes the estimator from the random-effects Poisson regression, and $\tilde{\beta}_{FE}$ denotes the estimator from the fixed-effects Poisson. If $T_H < \chi^2(\dim(\beta))$ at certain significance level, the null hypothesis that the unobserved individual characteristics (random-effects formulation) is not rejected. Obtaining estimates and variance-covariance matrices from both regressions, the computed $T_H = 98.39$, which is larger than $\chi^2(20)$ statistic at any significance level, therefore, the Hausman test strongly rejects the random-effects specification.

4.4 CONCLUSION

Since the Hausman specification test strongly rejects the random-effects specification, discussions in this section will focus on the results from the fixed-effects Poisson regression. The results from the fixed-effects Poisson regression suggest two major sources for missing entries in the labor data—labor market segmentation by age and sex, and intrahousehold allocation of labor resources. It is a problem that one cannot circumvent. Using only the complete observations will lead to biases in estimates and erroneous inferences, as discussed earlier. In fact, one should properly model individual labor market participation in a general equilibrium framework where risks and resources are shared at household level, and diversification in labor income will result with respect to changes in the socioeconomic status of the household. However, the intention of this essay is to understand the sources of missing observations, to explore possible biases in previous works that used the ICRISAT labor data, so that further empirical research using the data can be improved. The results from this essay are summarized into a bias adjustment statistic and brought into empirical tests for full insurance using the ICRISAT data in the third essay of the dissertation.

Table 4.1 Household and Member Annual Interview Times by Village

	Aurepalle	*Shirapur*	*Kanzara*
Household annual interview times	11.97	12.65	13.71
	(1.92)	(2.26)	(2.77)
Member annual interview times	10.22	10.97	11.54
	(3.02)	(2.85)	(4.07)

Note:

1. Standard deviations are reported in parentheses.
2. Household annual interview times are defined as the maximum number of visits a household receives regarding labor activities in a year. Member annual interview times are defined as the total counts of nonmissing data entries in the labor data in a year.

Table 4.2.a Fixed-Effects Poisson on Individual Annual Nonmissing Data Counts

Variable	*Coeff.*	*Std. Error (Robust)*
age	--	--
If age 15-44 male	0.24	(0.06)
If age 45-64 male	0.18	(0.06)
At least primary education	-0.11	(0.05)
# of 15-44 males in the household	--	--
# of 45-64 males in the household	--	--
# of members at least primary education (excluding the individual under study)	0.03	(0.00)
Household head at least primary education	--	--
Household head age in 45-64	--	--
Household (fitted) per capita income	-0.01	(0.003)
Year dummies (1975 omitted)		
1976	0.14	(0.02)
1977	0.14	(0.02)
1978	0.15	(0.03)
1979	0.12	(0.03)
1980	0.13	(0.04)
1981	0.15	(0.05)
1982	0.16	(0.05)
1983	0.16	(0.06)
1984	0.17	(0.07)
Exposure: total number of household visits		

Note:

1. Coefficients on time-invariant regressors such as constant, sex, family size, individual's main and secondary occupation, caste, household landholding class and household head sex are all absorbed in the individual-specific effects and are not identified.
2. Estimates presented are all at 95% confidence interval.
3. Reported are the actual coefficients, not the exponentiated ones that usually are interpreted as the incidence rate ratios.

Table 4.2.b Random-Effects Poisson on Individual Annual Nonmissing Data Counts

Variable	*Coeff.*	*Std. Err. (Robust)*
age	0.002	(0.000)
sex	-0.13	(0.03)
If age 15-44 male	0.17	(0.03)
If age 45-64 male	0.13	(0.03)
Relation with household head		
If direct family member	-0.06	(0.01)
At least primary education	-0.06	(0.01)
Main occupation in agriculture	0.03	(0.01)
Caste	--	--
Family size	-0.06	(0.02)
# of 15-44 males in the household	--	--
# of 45-64 males in the household	--	--
# of members at least primary education (excluding the individual under study)	0.02	(0.00)
Household head at least primary education	--	--
Household head sex (=1 if male)	0.08	(0.04)
Household head age in 45-64	--	--
Household (fitted) per capita income	-0.015	(0.003)
Household landholding class (Landless is omitted)		
Small farm	--	--
Medium farm	--	--
Large farm	--	--
Year dummies (1975 is omitted)		
1976	0.17	(0.02)
1977	0.16	(0.02)
1978	0.17	(0.02)
1979	0.11	(0.02)
1980	0.13	(0.02)
1981	0.13	(0.02)
1982	0.13	(0.02)
1983	0.13	(0.02)
1984	0.12	(0.02)
Constant	-0.22	(0.03)
Exposure: total number of household visits		

Note:

1. Estimates reported are all at 95% confidence interval.
2. Reported are the actual coefficients, not the exponentiated ones that usually are interpreted as the incidence rate ratios.
3. The random effects are assumed to be i.i.d. at individual level following a gamma distribution.

CHAPTER 5 TOWNSEND REVISITED

One of the most influential works in the literature of informal insurance is Townsend (1994). Using the longitudinal household survey data collected by the International Crops Research Institute for Semi-Arid Tropics (ICRISAT) from three villages in rural south India, Townsend (1994) tests how much protection the village-level informal mutual insurance provides against idiosyncratic risks by examining individual consumption.

5.1 TOWNSEND'S METHODOLOGY

Townsend (1994) starts by taking a village as a closed economy in which information flows and internal enforcement mechanisms are perfect. A representative individual derives utility from consumption and leisure. Preferences are constant absolute risk averse (CARA), uniform in the degree of risk aversion for all individuals, and additively separable in consumption and leisure. There are total S possible states of nature, uncorrelated across time. In each state, households receive idiosyncratic risks, although the exact nature of the risks is not assumed explicitly in Townsend's theoretical framework.

From the discussions in Chapter 3 of this dissertation, we know that by adopting additive separability in preferences, one implicitly imposes very strong restrictions on the substitution pattern of utility arguments. As a result, individual choice variables can be expressed in terms of the corresponding community-wide choice variables and a set of individual-specific factors. In the case with Townsend's, where individual consumption is the focus of theoretical and empirical interests, individual consumption can be fully expressed by village average consumption and individual characteristics.

A particular problem with the ICRISAT data is that consumption data were collected at household level instead of individual level. To incorporate this feature in the theoretical framework so that the predictions can be taken into estimations consistently, Townsend introduced "adult-equivalent" consumption and incomes. They are the age-sex adjusted average household consumption and incomes, with the adjustment factors obtained from external sources[1]. Similar to individual consumption, the equilibrium adult-equivalent consumption of a household is then expressed in terms of the village average adult-equivalent consumption and a set of individual-specific factors.

[1] The age-sex indices adopted in Townsend (1994) are based on a dietary survey by Ryan, Bidinger, Pushpamma, and Rao (1985): 1.0 for adult males, 0.9 for adult females, 0.94 for males13-18, 0.8 for females aged 13-18, 0.67 for children ages 7-12 regardless of gender, 0.52 for toddlers ages 1-3, and 0.05 for infants.

Under additive CARA preferences, village-level full insurance implies that conditioning on household-specific factors, the adult-equivalent consumption of a household comoves perfectly with the average adult-equivalent consumption of the village and nothing else. Therefore, full insurance can be tested by regressing the adult-equivalent consumption of a household on the average adult-equivalent consumption of the village and a set of "exogenous" variables such as household incomes, illnesses, and unemployment. In reduced forms, this means

$$(5.1)\ C^{*}_{it} = \alpha_i + \beta \overline{C}_t + \delta_i A_{it} + \varsigma_i X_{it} + \upsilon_{it},$$

where C^{*}_{it} is the adult-equivalent consumption of household i at date t, $\overline{C}_t$ is the average adult-equivalent consumption of the village at date t, A_{it} denotes demographic factors of the household, X_{it} denotes the set of exogenous variables, and υ_{it} is an i.i.d. disturbance term. Full insurance is rejected if any of the X_{it}'s is found to have significant impacts on the adult-equivalent consumption of a household.

Tests for full insurance in Townsend's are carried out in a series of univariate time-series and panel regressions. The dependent variable is $(C^{*}_{it} - \overline{C}_t)$, the mean difference of household i's adult-equivalent consumption from the village average adult-equivalent consumption. Right-hand-side explanatory variables include incomes, proportion of year sick, proportion of year unemployed, average village leisure, average village labor, family size, the number of adults in the households, and the number of children in the households.

In time series, regressions are run for each household with respect to each explanatory variable. There are around 40 household time series for each village, and at most 10 observations in each household time series. First-order serial correlation was rejected by the data. The results from time-series regressions fit the benchmark (5.1) remarkably well. Full insurance is rejected only for some 3 to 7 households in each village (see Table VI, Townsend, 1994).

In panel regressions, households from the same village are pooled to form a panel. There are 3 panels, and about 400 observations in each panel. One problem with the panel regressions such as (5.1) is that the average of the adult-equivalent consumption across households is close to the right-hand-side average consumption variable. This makes the coefficient on average adult-equivalent consumption of the village unity. Another problem is that there may exist measurement errors in the right-hand-side variables X_{it},

$$(5.2)\quad \hat{X}_{it} = X_{it} + v_{it},$$

and we observe $\hat{X}_{it}$ instead of the true X_{it}. To avoid the problem of perfect correlation and estimate (5.1) consistently, Townsend assumed first-order serial correlation in the disturbance terms and adopted a two-step approach. In the first step, the "within" estimates ς_w is obtained by:

$$(5.3)\ C^*_{it} - \overline{C}_t = \alpha_i + \delta_i A_{it} + \varsigma_w \hat{X}_{it} + \upsilon_{it} - \varsigma_w v_{it}.$$

In the second step, the estimates ς_Δ from a first-differenced version of (5.3) is obtained from

$$(5.4)\ \begin{aligned}(C^*_{it} - \overline{C}_t) - (C^*_{it-1} - \overline{C}_{t-1}) &= \alpha_i + \delta_i (A_{it} - A_{it-1}) + \varsigma_\Delta (\hat{X}_{it} - \hat{X}_{it-1}) \\ &+ (\upsilon_{it} - \upsilon_{it-1}) - \varsigma_\Delta (v_{it} - v_{it-1})\end{aligned}.$$

If there is no measurement error, the estimates on ς_w and ς_Δ do not differ significantly. Since $\varsigma_w - \varsigma_\Delta$ has an asymptotically normal distribution, measurement errors can tested by constructing a *t*-statistic from (5.3) and (5.4) for the null hypothesis that $H_0 : \varsigma_w - \varsigma_\Delta = 0$. If the null is rejected, a consistent estimate ς can be obtained according to

$$(5.5)\ \varsigma = \frac{2\varsigma_w \mathrm{Var}(\hat{X}) - \frac{T-1}{T}\varsigma_\Delta \mathrm{Var}(X_\Delta)}{2\mathrm{Var}(\hat{X}) - \frac{T-1}{T}\mathrm{Var}(X_\Delta)}$$ (see Townsend, 1994).

Townsend's estimates on ς_w and ς_Δ from the two-stepped regressions do not show significant difference. Regarding tests for full insurance, his results suggest that "...household consumption are not much influenced by ... sickness, unemployment, or other idiosyncratic shocks (page 539, Townsend, 1994)". Full insurance is nonetheless rejected because the coefficients on incomes are found to be significant, although the magnitudes of the estimates are notably small.

There are a number of issues in Townsend's that need special attention. Some of them are methodological problems, and some of them are problems in handling the data. One may naturally wonder how the results and inferences may change if we re-estimate Townsend's regressions with suitable corrections. In fact, as shown in the following, these problems are far from trivial. Table 5.1 summarizes the problems in Townsend's, my corrections, and my results in comparison with Townsend's. Tables 5.2 and onward provide detailed comparisons for each of my corrections.

5.2 DATA PROBLEMS IN TOWNSEND (1994)

The ICRISAT labor data are known to be noisy and awkward to use (see Kochar, 1995, 1999). Information on labor activities was collected and coded differently during 1975-1978 and 1979-1984. Between 1975 and 1978, interviewees' working hours and labor incomes were recorded by the type of labor activity engaged. Labor activities were classified into groups such as crop production, animal husbandry, handicrafts and trade, buildings, and repairs. A "partner code" is available and can be used to infer the employment status of an interviewee. Between 1979 and 1984, working hours and labor incomes were recorded by the nature of employment engaged. Labor activities were classified into three groups—farm production, government jobs, and non-farm-non-government jobs. It is possible to combine the information on labor utilization from the two parts of labor data through the partner code.

A small but nontrivial problem with the ICRISAT labor data is that the interviewees were coded slightly differently in the two parts of the survey. For example, an individual in household "A101 "during 1975-1978 normally was coded as "A10109". During 1979-1984, his or her code would become "A1019". When combining the information from these two parts, one needs to exert some special efforts to conform the identification codes.

Nonetheless, there are some nice consistencies in the labor data throughout the survey years. In particular, the reasons for not working since the last interview and the days of involuntary unemployment had been recorded consistently. These are of key interests to this dissertation.

5.2.1 SAMPLE SIZE AND PANEL REGRESSIONS

Only the 1979-1984 part of the ICRSAT labor data was used in Townsend's panel regressions. The reason is because the ICRISAT labor data "...are noisy and award to use..." (see page 546, Townsend, 1994). For example, when testing whether or not idiosyncratic health risks affect an individual's consumption, $(C^{*}_{it} - \overline{C}_{t})$ is regressed on the proportion of yearly sick in a univariate fashion in Townsend (1994). There are only 6 observations on the proportion of yearly sick because they are imputed from the 1979-1984 part of the labor data.

In Townsend's panel regressions, incomes are found to be significant, whereas the proportion of year sick and the proportion of involuntary unemployment are insignificant (see Table VIII, Townsend, 1994). The estimates are notably small and noisy. These small magnitudes, especially those on sickness and involuntary unemployment, lead to the aforementioned conclusion that "...household consumption are not much influenced by contemporaneous own...sickness, unemployment...shocks..." Knowing that the explanatory variables on labor utilization in Townsend's have only 6 years of data, a loss of 33 percent in sample size, one may wonder how valid these inferences are. Therefore, I revisit Townsend's panel regressions on sickness and unemployment with the complete data set. As in Towensend (1994), I assume measurement errors in sickness and unemployment, and estimate adult-equivalent all consumption and grain consumption in the two-stepped fashion. The estimates on ς_w, ς_Δ, and the consistent estimates are reported on Tables 5.2.a, 5.2.b, and 5.2.c along with Townsend's estimates (Table VIII panel a and panel b, Townsend, 1994).

Grain consumption is estimated separately because it is the key component of food and nutrition, therefore a crucial factor determining the basic wellbeing of an individual. Other parts of consumption such as large durables and luxurious goods are much more sensitive to changes in overall economic conditions. When insurance is perfect, all consumption components are smoothed, therefore the coefficients on the proportion of year sick and the proportion of involuntary unemployment from both regressions are not significantly different from zero. When full insurance is rejected in the all consumption regressions but not in the grain consumption regressions, we can infer that insurance is not perfect, yet households' basic wellbeing is reasonably protected. When full insurance is rejected in both regressions, we can infer that households' food and nutrition are affected. If we knew more about the actual health status of the study individuals, we would be able to infer how imperfect insurance affects health and labor productivity. Clearly, the last case is the most severe in terms of welfare losses.

Tables 5.2.a and 5.2.b provide my estimates of ς_w and ς_Δ on the proportion of year sick and the proportion of involuntary unemployment. Similar to Townsend's, with one exception, none of my estimates on the proportion of year sick are statistically significant. The t-statistics from Table 5.2.c suggest that there is little systematic difference between these two estimators. If the proportion of year sick truly reflects the occurrence and severity of the idiosyncratic health risks the ICRISAT households face, the results suggest that a household's consumption and basic wellbeing are well protected against idiosyncratic health risks. From this perspective, it seems that the underlying village mutual insurance can provide reasonably good risk protections resembling that of a formal health insurance.

Unlike Townsend's, my estimates on the proportion of involuntary unemployment are positive and significant in most regressions. For the village of Shirapur, the coefficients on the proportion of involuntary unemployment in all consumption and grain consumption regressions are all positive and significant. For the village of Aurepalle, three out of four estimates are positive and significant. The case with Kanzara seems weaker, only one out of four estimates are positive and significant. The t-statistics presented in Table 5.2.c suggest *some* evidence of measurement errors in the proportion of involuntary unemployment. The problem seems particularly severe in the grain consumption regressions. After correcting for measurement errors, the consistent estimates from grain consumption regressions are much higher than those from the within estimations and first-difference estimations.

Combining information from Tables 5.2.a-5.2.c, it seems that involuntary unemployment is positively related with consumption. One may find it difficult to interpret and even counter-intuitive. A reverse causal relationship may seem more plausible here. An individual's reservation wage usually varies with his or her socioeconomic status. If we use consumption as a proxy for the socioeconomic status of the individual, as consumption increases, so does reservation wage. Individuals with higher reservation wage are more likely to be (involuntarily) unemployed in certain segments of the labor market. If this is the case, the coefficients on the proportion of involuntary unemployment shall not be taken as evidence against full insurance.

The results from grain consumption regressions seem to support this view. Notice that the estimates are uniformly positive and significant for all villages, a much stronger case than the all consumption regressions. As discussed in Chapter 2 of this dissertation, improved food consumption and nutrition can increase labor productivity, at least for those who have very low incomes. Some empirical works have shown the existence of "nutrition-based efficiency wage" in the labor markets in developing countries (see Dasgupta, 1997, Pitt et al, 1990, and Deolalikar, 1988). Labor markets are segmented by the nutritional status of participants, and it is "efficient" for employers to pay higher wage rate to those who have better nutrition. When the market segments for better nutrition do no exist, involuntary unemployment emerges. Again, in this case, the coefficients on the proportion of involuntary unemployment shall not be taken as evidence against full insurance.

The nature of involuntary unemployment and the nutrition-based efficiency wage hypothesis could be explored further if there is information on individuals' nutritional status and health outcomes. However, the current data set for the three study villages do not contain such information.

5.2.2 MISSING OBSERVATIONS AND SAMPLE SELECTION PROBLEMS IN TOWNSEND (1994)

In Chapter 4 of this dissertation, the problem of missing observations in the ICRISAT labor data is discussed in great detail. 12.6 percent of the observations are missing, and the pattern of missing observations is not random. With discrete estimations, Chapter 4 shows that observations on labor utilization are missing due to three reasons— the efforts made by the investigator, the characteristics and qualifications of market participants, and the socioeconomic status and intrahousehold allocations of the households. In particular, conditioning on other factors, individuals from lower income families are more "visible" in the village labor markets. This problem has never been discussed in any of the existing empirical works using the ICRISAT labor data, including Townsend's.

The proportions of year sick and involuntary unemployment can only be derived from complete observations. If missing observations are due to self-selection, those who have lower incomes are included in the regressions. When estimating the proportions of year sick and involuntary unemployment, the inferences are only pertinent to those that are included in the regressions, because there is no way for us to know whether or not the excluded individual/observations had lost work due to health reasons or had been involuntarily unemployed since the last interview. Extending the inferences to the whole sample or even the whole village in this case is definitely not appropriate. To understand how serious the biases are in Townsend's regressions in this regard and to obtain valid inferences for the whole sample, I incorporate the results from Chapter 4 of this dissertation and revisit Townsend's panel regressions. Since random-effects Poisson specifications in Chapter 4 are rejected by Hausman tests, the following focuses only on the results from fixed-effects Poisson regressions.

Based on the results from Chapter 4, we can obtain a predicted interview frequency for each individual. Denote $\hat{N}_i$ as the predicted interview frequency for individual *i*. $\hat{N}_i$ is given by

$$(5.6)\quad \hat{N}_i = F\left(\hat{\theta}_i, Z_i\right),$$

where $F(\cdot)$ is the c.d.f of fixed-effects Poisson, $\hat{\theta}_i$ denotes the vector of estimated coefficients in Table 4.2.a, and $\hat{Z}_i$ is the vector of explanatory variables in Table 4.2.a. Denote S_i as the selection factor. It is the ratio between predicted interview frequencies and the actual interview frequencies an individual receives. That is,

$$(5.7)\ S_i = \hat{N}_i / N_i.$$

Define the selection-adjusted proportion of year sick, $\hat{T}_i$, as

$$(5.8)\ \hat{T}_i = \hat{S}_i \times T_i,$$

where T_i is the recorded proportion of year sick. Define the selection-adjusted proportion of involuntary unemployment, $\hat{U}_i$, as

$$(5.9)\ \hat{U}_i = \hat{S}_i \times U_i,$$

where U_i is the recorded proportion of involuntary unemployment.

In the following, I re-estimate adult-equivalent all consumption and grain consumption in the two-stepped fashion as before, but the regressors are all selection-adjusted proportions according to (5.8) and (5.9). The results are presented in Tables 5.3.a, 5.3.b, and 5.3.c.

We can see that the results are very similar to those presented in Tables 5.2.a, 5.2.b, and 5.2.c. The small t-statistics suggest that there is little evidence of measurement errors in the selection-adjusted proportion of year sick. Unlike those from Tables 5.2.a, 5.2.b, and

5.2.c in which none of the estimates on the proportion of year sick is significant, here we see that it is negative and significant in the all consumption regressions for the village of Aurepalle after adjusting for selection problems. However, the estimates from the corresponding grain consumption regressions are not statistically significant. We can infer that the consumption of individuals in Aurepalle are not well protected against idiosyncratic health risks. On one hand, an individual's consumption tends to decrease as the loss of work due to illnesses and injuries increases. On the other hand, imperfect insurance seems to have very little impacts on an individual's basic wellbeing (the estimates on the selection-adjusted proportion of year sick from grain consumption regressions are not statistically significant.)

We can further infer how much income an individual would like to give up in order to completely eliminate the variations in consumption based on the derivations in Chapter 1 of this dissertation.

5.2.3 IMPUTATION OF LEISURE

In the panel regressions, Townsend also included village average leisure and village average labor supply to test full insurance. Townsend argued that the average village supply "...is a proxy for average leisure, one which avoids the problem of measuring time endowment...(page 547, Townsend, 1994)."

Ideally, the imputation of leisure can be done by subtracting sick days, unemployed days, days not working due to social or personal reasons, time spent on domestic work and home production from a total time endowment (Townsend used 312 days as the non-holiday time endowment for each year). In 1975-1978, hours spent on domestic work and home production such as food and fuel gathering were recorded in the labor data, so that it is possible for us to impute the hours of leisure of each interviewee. In 1979-1984, however, information on domestic work and home production was not collected, and this is the very part of the labor data used by Townsend.

Since there exist alternative tests for full insurance, it is not necessary to include village average leisure in the regressions, because it is impossible to impute leisure safely from the complete labor data.

5.3 PRICE INDEX AND REGIONAL CHARACTERISTICS

Consumption and income measures in Townsend (1994) are deflated by the 1975 all-India prices. The problem of using a national price index is that certain regional and village-specific characteristics may be represented disproportionally. For example, the study households are very poor. In fact, 75 percent of them live below India's absolute poverty line (Walker and Ryan, 1990), and their incomes are mainly derived from agricultural productions and labor-related activities (see Chapter 2 of this dissertation). There were formal health care system, no formal financial services, and no modern infrastructure of any kind during the survey years. All of these suggest that labor earnings and agricultural productions deserve much higher weights when their nominal values are deflated. An all-India price index under-represents the importance of agricultural activities, and tends to ignore regional characteristics.

The problem of under-representation of agricultural activities is not serious in Townsend's time-series regressions, because both the dependent variable and explanatory variables are adjusted by the same all-India price index for each year, regressions are all carried out in univariate fashions, and the first-order serial correlation is rejected. However, in Townsend's panel regressions where households from the same village are pooled together, the estimates on incomes reflect the average response for all households in all ten years. Deflating the nominal values is no longer a simple matter of symmetric scaling.

Due to these concerns, I revisit Townsend's panel regressions again with the village-specific price indices from Walker and Ryan (1990). As in Townsend's, only the later labor data is used in estimations, so that the effects can be shown separately from the effects resulting from increased sample size and sample selection adjustments. The results are reported in Table 5.4.a, 5.4.b, and 5.4.c.

Compared with Townsend's, I find that measurement errors are particularly severe in the proportion of year sick and the proportion of involuntary unemployment (see Table 5.4.c), which seems to be compatible with the findings in the previous section and in Chapter 4 of this dissertation. If one only includes the complete observations in regressions, the unobserved factors determining self-selection are ignored. From this perspective, we can well say that the proportion of year sick and the proportion of involuntary unemployment are observed with errors.

Townsend found that "...income does matter in the determination of consumption. Still one may be surprised at the small magnitude of many of these estimates." (page 578, Townsend, 1994) Using the village-specific prices, my estimates on crop incomes and labor incomes are between 3 to 10 times as high as Townsend's and much more significant for all three villages. For the villages of Shirapur and Kanzara, my estimates on the profits from trade and handicrafts and the profits from animal husbandry are up to 10 times as high as Townsend's. All of these suggest that the importance of incomes from labor-intensive production activities to the study household is seriously under-represented by using a national price index. Incomes actually have a much larger impact on consumption.

Similar to the findings in the previous section, we can see here that the proportion of year sick tends to have a negative and significant impact on consumption when village-specific prices are used. In areas such as the ICRISAT villages where incomes are mainly derived from labor activities and where there is no formal health care or health insurance, human illnesses and injuries not only incur monetary loss associated with seeking treatment, but also the loss of income due to the loss of working time. If labor-related incomes are found to have positive impacts on consumption, it is not surprising to find that the proportion of year sick to have a negative impact on consumption, because the longer sickness, the more losses in income. When using national price indices, the importance of the loss of labor income due to illnesses and injuries is under-represented, which may help explain why Townsend's estimates on the proportion of year sick are all very small.

Table 5.1 Comparison of Yi's and Townsend's Results

Townsend		*Yi*	
Problem	***Results***	***Enhancement***	***Results***
Sample size: 1979-84 → 6 observations in each panel regression	Proportion of yearly involuntary unemployment: insignificant	Sample size: 1975-84 10 observations in panel regressions	Proportion of yearly involuntary unemployment: most are + & significant
Missing observations: not addressed → imputation of yearly involuntary unemployment and sickness based on complete observations	Proportion of yearly involuntary unemployment & sickness insignificant	Missing observations: sample selection incorporated	Proportion of yearly sickness: some are - & significant Proportion of yearly involuntary unemployment: most are + & significant
Price index: 1975 all-India prices → underrepresentation of agriculture-related activities for the HHs	Small estimates on agricultural incomes. Small & insignificant estimates on yearly involuntary unemployment & sickness	Price index: 1980 village-specific prices	Larger estimates on agricultural incomes. Larger & more significant estimates on yearly involuntary unemployment & sickness

Table 5.2.a Yi vs. Townsend: Panel Within Estimates (10-Year Data)

		Townsend: 1979-1984 data		Yi: 1975-1984 data	
Population	Variable	All Consumption	Grain Consumption	All Consumption	Grain Consumption
Aurepalle	Proportion of Year Sick	1462.02	39.353	-230.61†	-22.722
		(1469.02)	(1477.63)	(96.485)	(61.506)
	Proportion of Involuntary Unemploy.	NA	NA	-63.089	344.46*
				(184.74)	(66.319)
Shirapur	Proportion of Year Sick	NA	NA	77.472	18.022
				(92.076)	(22.481)
	Proportion of Involuntary Unemploy.	NA	NA	247.30†	206.40*
				(127.07)	(27.083)
Kanzara	Proportion of Year Sick	-0.94	59.170	-38.745	-35.362
		(325.22)	(269.53)	(76.434)	(23.537)
	Proportion of Involuntary Unemploy.	1192.96	789.276	-32.926	251.52*
		(752.48)	(862.84)	(303.41)	(69.822)

Note: Standard errors are reported in parentheses. * at 95% confidence level. † at 90% significance level.

Table 5.2.b Yi vs. Townsend: Panel First-Difference Estimates (10-Year Data)

		Townsend: 1979-1984 data		Yi: 1975-1984 data	
Population	Variable	All Consumption	Grain Consumption	All Consumption	Grain Consumption
Aurepalle	Proportion of Year Sick	493.741	171.615	-131.905	3.2417
		(1852.4)	(1812.26)	(127.73)	(58.233)
	Proportion of Involuntary Unemploy.	NA	NA	455.908*	247.175*
				(182.74)	(89.241)
Shirapur	Proportion of Year Sick	NA	NA	25.256	-4.4483
				(49.632)	(19.805)
	Proportion of Involuntary Unemploy.	NA	NA	132.34†	99.380*
				(71.805)	(28.601)
Kanzara	Proportion of Year Sick	269.242	172.770	6.1872	-42.302†
		(280.49)	(298.77)	(64.453)	(31.524)
	Proportion of Involuntary Unemploy.	60.6126	456.984	120.300	56.161
		(534.058)	(792.377)	(187.93)	(91.983)

Note: Standard errors are reported in parentheses. * at 95% confidence level. † at 90% significance level.

Table 5.2.c Yi vs. Townsend: Measurement Error and Consistent Estimates (10-Year Data)

Population	Variable	Townsend: 1979-1984 data All Consumption	Townsend: 1979-1984 data Grain Consumption	Yi: 1975-1984 data All Consumption	Yi: 1975-1984 data Grain Consumption
Aurepalle	Proportion of Year Sick	[0.210]	[-0.094]	[0.7874]	[0.4243]
	Proportion of Involuntary Unemploy.	NA	NA	-319.758 [-3.8262]	[1.4663]
Shirapur	Proportion of Year Sick	NA	NA	[0.8544]	[0.9998]
	Proportion of Involuntary Unemploy.	NA	NA	[1.5427]	288.227 [3.9522]
Kanzara	Proportion of Year Sick	[-0.200]	[-0.200]	[0.7550]	[0.3000]
	Proportion of Involuntary Unemploy.	[0.259]	[0.259]	[0.8660]	351.970 [2.7982]

Note: Reported in squared brackets are the t-statistic testing $\varsigma_w = \varsigma_\Delta$. When rejecting $\varsigma_w = \varsigma_\Delta$, the consistent estimate is constructed according to (5.5).

Table 5.3.a Yi vs. Townsend: Panel Within Estimates
(Sample Selection Adjustment)

		Townsend		Yi	
		All	Grain	All	Grain
Population	Variable	Consumption	Consumption	Consumption	Consumption
Aurepalle	Proportion of Year	1462.02	329.640	-216.724[†]	-49.674
	Sick	(1469.02)	(772.506)	(119.062)	(58.625)
	Proportion of	NA	NA	-129.398	292.469*
	Involuntary Unemploy.			(139.097)	(68.154)
Shirapur	Proportion of Year	NA	NA	72.625	14.714
	Sick			(63.127)	(23.162)
	Proportion of	NA	NA	206.601*	195.726*
	Involuntary Unemploy.			(74.795)	(27.195)
Kanzara	Proportion of Year	-0.94	34.118	-54.503	-39.037
	Sick	(325.22)	(166.85)	(64.704)	(25.539)
	Proportion of	1192.96	250.140	-43.510	240.626*
	Involuntary Unemploy.	(752.48)	(388.82)	(178.419)	(70.259)

Note: Standard errors are reported in parentheses. * at 95% confidence level. [†] at 90% significance level.

Table 5.3.b Yi vs. Townsend: Panel First-Difference Estimates
(Sample Selection Adjustment)

		Townsend		Yi	
		All	Grain	All	Grain
Population	Variable	Consumption	Consumption	Consumption	Consumption
Aurepalle	Proportion of Year	493.741	329.64	-216.155*	-8.098
	Sick	(1852.40)	(772.51)	(105.785)	(48.267)
	Proportion of	NA	NA	358.690*	170.779*
	Involuntary Unemploy.			(173.743)	(79.185)
Shirapur	Proportion of Year	NA	NA	50.640	0.429
	Sick			(41.499)	(16.764)
	Proportion of	NA	NA	111.084	96.147*
	Involuntary Unemploy.			(71.648)	(28.537)
Kanzara	Proportion of Year	269.242	90.563	24.960	-33.843
	Sick	(280.49)	(204.62)	(67.186)	(32.876)
	Proportion of	60.6126	149.675	84.755	41.760
	Involuntary Unemploy.	(534.058)	(388.463)	(188.817)	(92.415)

Note: Standard errors are reported in parentheses. * are significant at the 95% confidence level.

Table 5.3.c Yi vs. Townsend: Measurement Errors and Consistent Panel Estimates (Sample Selection Adjustment)

		Townsend		Yi	
		All	Grain	All	Grain
Population	Variable	Consumption	Consumption	Consumption	Consumption
Aurepalle	Proportion. of Year				
	Sick	[0.210]	[0.046]	[0.000]	[0.7071]
	Proportion of	NA	NA	-434.032	368.4202
	Involuntary Unemp.			[3.5085]	[1.7861]
Shirapur	Proportion of Year	NA	NA		
	Sick			[0.3464]	[0.6164]
	Proportion of	NA		281.1553	273.4508
	Involuntary Unemp.		[0.250]	[1.2691]	[3.6620]
Kanzara	Proportion of Year		NA		
	Sick	[0.443]		[1.2288]	[0.2000]
	Proportion of				343.3195
	Involuntary Unemp.	[0.910]	[0.169]	[0.7211]	[2.8302]

Note: Reported in squared brackets are the t-statistics testing $\varsigma_w = \varsigma_\Delta$. When rejecting $\varsigma_w = \varsigma_\Delta$, the consistent estimate is constructed according to (5.5).

Table 5.4.a Yi vs. Townsend: Panel Within Estimates (1980 Prices)

		Townsend		Yi	
Population	Variable	All Consumption	Grain Consumption	All Consumption	Grain Consumption
Aurepalle	All Income	0.0772*	0.0474*	0.0479*	0.0142*
		(0.0221)	(0.0159)	(0.0097)	(0.0054)
	Crop Income	-0.0150	0.0238	0.2310*	0.0176
		(0.0312)	(0.024)	(0.0512)	(0.0214)
	Labor Income	0.0401	0.0591	-0.0088	0.2410*
		(0.0647)	(0.0464)	(0.1412)	(0.0926)
	Profit from Trade and Handicraft	0.2363*	0.1241*	0.0305*	0.0241*
		(0.0352)	(0.0260)	(0.0035)	(0.0033)
	Profit from Animal Husbandry	0.0485*	-0.1539*	0.0305*	-0.0068
		(0.0676)	(0.0478)	(0.0035)	(0.0389)
	Proportion of Year Sick	1462.02	329.64	-230.61*	-22.600
		(1469.77)	(772.506)	(96.485)	(50.656)
	Proportion of Involuntary Unemployment	NA	NA	-63.089	324.87*
				(184.74)	(100.67)
	# Household Members	-45.778*	-33.571*	-7.5912	-26.556*
		(6.3693)	(4.5549)	(18.175)	(4.9266)
	# Adults	-30.459*	-18.987*	17.673	-22.580*
		(9.7187)	(6.4754)	(17.210)	(5.6464)
	# Children	-47.880*	-28.557	-50.442*	-33.372*
		(11.9590)	(7.9957)	(19.751)	(8.1244)
Shirapur	All Income	0.1169*	0.0605*	0.0541*	0.0181*
		(0.0277)	(0.0129)	(0.0106)	(0.0033)
	Crop Income	0.0825*	0.0463*	0.1858*	0.0593*
		(0.0373)	(0.0175)	(0.0311)	(0.0113)
	Labor Income	0.1127*	0.1022*	0.1391*	0.0733*
		(0.0740)	(0.0345)	(0.0993)	(0.0376)
	Profit from Trade and Handicraft	0.00291	0.00447	0.0419*	0.0196*
		(0.0671)	(0.0315)	(0.0021)	(0.0007)
	Profit from Animal Husbandry	0.5014*	0.0937*	0.5242*	0.1341*
		(0.0789)	(0.0389)	(0.0938)	(0.0289)
	Proportion of Year Sick	NA	NA	77.472	17.972
				(92.076)	(27.667)
	Proportion of Involuntary Unemployment	NA	NA	247.30†	205.62*
				(127.07)	(39.486)
	# Household Members	-25.507*	-15.955*	-19.950*	-7.8337*
		(5.5078)	(2.5296)	(5.7861)	(1.7779)
	# Adults	-52.448*	-23.177*	-14.764	-6.7740
		(8.8253)	(4.2859)	(12.765)	(4.5106)
	# Children	-22.154*	-22.833*	-35.475*	-13.067*
		(10.221)	(4.7524)	(8.3211)	(2.9621)

Kanzara	All Income	-0.0073	0.0725*	0.1008*	0.0258†
		(0.0219)	(0.0122)	(0.0291)	(0.0150)
	Crop Income	0.0513*	0.0596*	0.1744*	0.0326
		(0.0286)	(0.0165)	(0.0525)	(0.0232)
	Labor Income	0.0198	0.0721*	0.2251	0.1515*
		(0.0406)	(0.0279)	(0.1891)	(0.0860)
	Profit from Trade and	0.1347	0.2794*	-0.0246	-0.0240†
	Handicraft	(0.0895)	(0.0569)	(0.0459)	(0.0132)
	Profit from Animal	0.0672	0.1132*	0.7700†	0.3053
	Husbandry	(0.0606)	(0.0476)	(0.4032)	(0.2395)
	Proportion of Year	-0.94	34.118	-38.745	-35.189
	Sick	(325.22)	(166.85)	(76.434)	(23.129)
	Proportion of	1192.96	250.140	-32.926	250.39*
	Involuntary	(752.48)	(388.82)	(303.41)	(105.59)
	Unemployment				
	# Household Members	-21.355*	-6.1817	-21.677*	-9.2096*
		(6.7116)	(3.9607)	(4.9129)	(2.8886)
	# Adults	-25.534*	3.9391	-22.241	-11.814*
		(9.3840)	(5.6423)	(11.034)	(3.3285)
	# Children	-12.7078	-11.8080	-36.334*	-12.492
		(12.866)	(7.6255)	(10.057)	(3.7276)

Note: Standard errors are reported in parentheses. * are significant at the 95% confidence interval. † are significant at the 90% confidence interval.

Table 5.4.b Yi vs. Townsend: Panel First-Difference Estimates (1980 Prices)

		Townsend		Yi	
Population	Variable	All Consumption	Grain Consumption	All Consumption	Grain Consumption
Aurepalle	All Income	0.0469	0.0289*	0.0372*	0.0147*
		(0.0236)	(0.0151)	(0.0064)	(0.0045)
	Crop Income	-0.0380	-0.0066	0.1024*	0.0229*
		(0.0299)	(0.0191)	(0.0271)	(0.0092)
	Labor Income	0.2597*	0.2335*	0.3000*	0.1850*
		(0.0830)	(0.0522)	(0.0880)	(0.0802)
	Profit from Trade and Handicraft	0.1495*	0.0430*	0.0280*	0.0088
		(0.0389)	(0.0252)	(0.0018)	(0.0057)
	Profit from Animal Husbandry	-0.0276	-0.0081	0.0149	0.0140
		(0.0689)	(0.0439)	(0.0735)	(0.0178)
	Proportion of Year Sick	493.741	329.64	-146.01	-143.82
		(1852.40)	(772.51)	(99.966)	(60.821)
	Proportion of Involuntary Unemployment	NA	NA	276.14†	206.87*
				(150.47)	(82.090)
Shirapur	All Income	-0.0592*	0.0233	0.0136	0.0059*
		(0.0236)	(0.0142)	(0.0057)	(0.0017)
	Crop Income	0.0352	0.0172	0.0412*	0.0208*
		(0.0301)	(0.0181)	(0.0184)	(0.0071)
	Labor Income	0.1925*	0.1456*	0.1463*	0.0672*
		(0.06550)	(0.0390)	(0.0452)	(0.0162)
	Profit from Trade and Handicraft	-0.1091	-0.0773	0.0171*	0.0065*
		(0.0757)	(0.0453)	(0.0006)	(0.0019)
	Profit from Animal Husbandry	-0.1994*	0.0118	0.1775*	0.4057*
		(0.0693)	(0.0420)	(0.0690)	(0.0157)
	Proportion of Year Sick	NA	NA	-4.7416	-13.444
				(118.06)	(31.932)
	Proportion of Involuntary Unemployment	NA	NA	224.25*	136.08*
				(104.08)	(33.735)
Kanzara	All Income	0.1233*	0.0697*	0.0299*	0.0124*
		(0.0227)	(0.0152)	(0.0080)	(0.0053)
	Crop Income	0.0677*	0.0313*	0.0547*	0.0152*
		(0.0308)	(0.0204)	(0.0194)	(0.0068)
	Labor Income	0.1003*	0.0721*	0.1411†	0.0769
		(0.0422)	(0.0279)	(0.0768)	(0.0678)
	Profit from Trade and Handicraft	0.4057*	0.2794*	0.0449*	-0.0025
		(0.0863)	(0.0569)	(0.0074)	(0.0542)
	Profit from Animal Husbandry	0.2252*	0.1132*	0.1752†	0.2170
		(0.0715)	(0.0476)	(0.1005)	(0.1745)
	Proportion of Year Sick	269.242	90.563	16.490	-26.495†
		(280.49)	(204.62)	(55.936)	(15.272)
	Proportion of Involuntary Unemployment	60.6126	149.675	235.88†	186.17*
		(534.058)	(388.463)	(117.95)	(61.600)

Note: Standard errors are reported in parentheses. * are significant at the 95% confidence interval. † are significant at the 90% confidence interval.

Table 5.4.c Yi vs. Townsend: Measurement Errors and Consistent Estimates (1980 Prices)

		Townsend		Yi	
Population	Variable	All Consumption	Grain Consumption	All Consumption	Grain Consumption
Aurepalle	All Income				
		[0.768]	[0.599]	[0.0165]	[0.0065]
	Crop Income				
		[0.380]	[0.7136]	[0.1193]	[0.0392]
	Labor Income				
		[-1.543]	[-1.716]	[-0.3097]	[-0.0365]
	Profit from Trade and		[1.538]		
	Handicraft	[1.197]		[-0.0079]	[0.0053]
	Profit from Animal				
	Husbandry	[-0.116]	[-1.622]	[0.2585]	[0.0660]
	Proportion of Year		NA	-559.44	-344.20
	Sick	[0.210]		[-12.326]	[-56.023]
	Proportion of	NA		3616.03	-2066.1
	Involuntary		[-0.511]	[252.46]	[9.9084]
	Unemployment				
Shirapur	All Income				
		[1.290]	[1.676]	[0.0402]	[0.0234]
	Crop Income				
		[0.690]	[0.818]	[0.1316]	[0.0781]
	Labor Income				
		[-0.271]	[-0.497]	[-0.0006]	[0.0252]
	Profit from Trade and		0.3276		
	Handicraft	[0.742]	[2.074]	[0.0250]	[0.0214]
	Profit from Animal	1.4678			
	Husbandry	[2.193]	[1.204]	[0.3583]	[0.2083]
	Proportion of Year	NA	NA	-1043.22	-497.84
	Sick			[182.32]	[33.304]
	Proportion of	NA	NA	-28.831	-481.39
	Involuntary			[91.298]	[53.190]
	Unemployment				
Kanzara	All Income	0.2177			
		[-3.177]	[0.120]	[0.0610]	[0.0205]
	Crop Income	-0.2545			
		[-2.355]	[0.935]	[0.1056]	[0.0345]
	Labor Income				
		[-1.058]	[-0.278]	[0.0015]	[0.0574]
	Profit from Trade and				
	Handicraft	[-1.312]	[-1.562]	[-0.0538]	[-0.0169]
	Profit from Animal				
	Husbandry	[-1.387]	[-0.243]	[0.3785]	[0.0816]
	Proportion of Year			-382.71	120.86
	Sick	[-0.200]	[-0.250]	[-52.570]	[-20.648]
	Proportion of			7937.48	-2651.10
	Involuntary	[0.259]	[0.169]	[-443.15]	[-24.638]
	Unemployment				

Note: Standard errors are reported in parentheses. * are significant at the 95% confidence interval. † are significant at the 90% confidence interval.

CHAPTER 6 HEALTH RISKS AND FULL INSURANCE

6.1 HEALTH RISKS IN THE ICRISAT VILLAGES

In the ICRISAT study villages, poor sanitation, hygiene, and public health infrastructure are most likely responsible for health problems. According to Walker and Ryan (1990), sanitation was inadequate in all the study villages. Only 15 percent of the households satisfied the good conditions of household cleanliness. None of the villages has a trained physician. The nearest public health care center ranged from 4 km to 13 km, a distance to be managed with quite some time and efforts. Nutritional status and caloric intake among the study households were also poor relative to the national averages in India. According to a separate survey in 1976 conducted at the ICRISAT villages, the mean intake for all age-sex groups was less than half the Indian standards for beta-carotene and ascorbic acid, less than three-quarters of the Indian standards for calcium and riboflavin, and below the Indian standards for energy. Caloric deficiencies were most severe among young children. As age increased, caloric deficiencies only declined slightly. Clinical symptoms of vitamin A and vitamin B-complex deficiencies were evident in preschool children in all the villages. About 8 percent of individuals fell below 70 percent of their age-sex weight-for-height standards.

Because of these conditions, sickness is no stranger to the villages. According to a 52-week survey conducted in Dokur, the weekly sickness incidence rate (the proportion of individuals with morbid signs) ranges from 14 to 25 percent of the sample (see Walker and Ryan, 1990, and Bidinger, Nag, and Babu, 1986).

On the other hand, the reported rate of sickness leave from work since the last survey (roughly three to four weeks before) averaged only about 2.5 percent. Positive medical spending in cash or in kind was almost none. This seems to be rather inconsistent with the aforementioned sickness incidence rate. It seems as though the individuals did not seek enough medical treatment or take as much sickness leave from work in response to the high sickness incidences. A natural hypothesis for the disproportional responses in labor supply with respect to sickness incidences is that the individuals were not well protected against health risks. For example, lacking income transfers through mutual insurance upon the incidence of sickness, an individual living close to destitute cannot afford to take sickness leave to for recoveries.

The disproportional responses in sickness leave are in contradiction to the findings from the famous work of Townsend. Using the ICIRSAT data, Townsend (1994) finds that households are relatively well protected against idiosyncratic health risks (see Townsend, 1994). With a closer look, we know that the data set does not contain any information on individuals' true health status, or the incidences and severity of illnesses and injuries. Only the incidence of sickness leave from the labor market was recorded in the data, which was taken as a proxy for health risks by Townsend. Given the fact that the actual sickness incidences are much higher than the reported rate of sickness leave, it is not surprising that the proportion of year sick does not show any statistical significance in Townsend's regressions. Sickness leave, as argued earlier, is a choice variable under an individual's discretion, because it incurs losses in labor income and affects consumption if insurance is imperfect. The proportion of year sick cannot be used directly to infer the effectiveness of the village mutual insurance on idiosyncratic health risks. To make valid inferences, we shall model individuals' decision on taking sickness leave properly.

In Chapter 3 of this dissertation, individuals are assumed to derive utility from consumption, health, and leisure. An individual's health is reduced from the natural level by i.i.d. risks, and can be recovered by taking sickness leave. The theoretical predictions on consumption, health, and sickness leave derived under three preference specifications are very intuitive and can be taken into estimations very easily. Therefore, the empirical estimations in this chapter are based on the theoretical framework developed in Chapter 3 of this dissertation.

6.2 CONSUMPTION AND HEALTH UNDER FULL INSURANCE

6.2.1 TESTS FOR FULL INSURANCE UNDER DIFFERENT PREFERENCE SPECIFICATIONS

Health is one of the most important aspects of an individual's wellbeing—it is a highly valued asset by itself, a prerequisite for enjoying consumption and leisure, and a prerequisite for labor earnings. In developing countries, especially in the rural areas, good health is of particular importance because production technologies are labor-intensive. Good health does not only serve as a basis for deriving labor income, but also reduces the monetary costs associated with seeking medical treatment, and the opportunity costs resulting from the loss of work due to illnesses and injuries.

In addition to good health, there are other goals in life, such as consumption. From Chapter 3 of this dissertation, we know that there is no tradeoff between health and consumption if an individual ismeanings under different preference specifications. Under separable preferences (additive CARA and multiplicative CRRA), conditioning on individual-specific factors and the village aggregate variables, full insurance implies that an individual's consumption is not affected by his or her health status and health risks.

The relative ratio between health and consumption under full insurance can be expressed as

$$\text{(6.1)} \quad \frac{H_{is}}{C_{is}} = f\left(\theta_i, w_{is}, \overline{K}_s\right),$$

where parameter θ_i represents individual i's characteristics such as risk attitude and welfare weight, w_{is} represents individual i's market wage rate, and $\overline{K}_s$ represents the village aggregate economic resources. By "no tradeoff", we mean that the ratio between health and consumption is not affected by the idiosyncratic health risks, or, consumption is not excessively sensitive to changes in health, other things equal.

Under general nonseparable preferences, the substitutability between consumption and health varies at different levels of consumption and health. Due to nonseparability between, the determination of an individual's consumption under full insurance necessarily involves health. The relative ratio between health and consumption cannot be easily derived, but the relationship between the percentage changes of health and consumption can be derived from

$$\text{(B.14)} \quad \xi_{CH}\Delta H \cong \xi_{HC}\Delta C - \Delta w.$$

Following the notations in (6.1), (B.14) implies that there exists an implicit function relating the percentage change of health and the percentage change of consumption as follows,

$$\text{(6.2)} \quad \frac{\Delta H_{is}}{\Delta C_{is}} = g\left(\frac{\xi_{HC}}{\xi_{CH}}, w_{is}\right),$$

a ratio that depends on the substitutability between health and consumption (the elasticity terms), and the market wage rates. Notice that it also changes with the levels of consumption and health because the elasticitities varies at different consumption and health levels. $\frac{\Delta H}{\Delta C}$ also has the interpretation of elasticity of health with respect to consumption. By "no tradeoff", we mean that health does not change excessively with consumption variations more than the elasticity.

Tests for full insurance for all three preference specifications can be carried out under a unified framework such as

$$(6.3)\quad \begin{cases} C_{it} - \overline{C}_t = \alpha_i^1 + \gamma_i^1 T_{it} + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} - \overline{T}_t = \alpha_i^2 + \gamma_i^2 C_{it} + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + \eta_i w_{it} + v_{it} \end{cases}$$

(see Chapter 3 of this dissertation for more discussions). The γ's reflect the degree of nonseparability between consumption and health. The ς's on the "exogenous" variables measure how the socioeconomic status of an individual such as incomes, wealth, and family demographic characteristics affect his or her consumption and health. Under imperfect insurance, these factors will restrict the scope and coverage of mutual insurance an individual receives. For example, when insurance is imperfect, a wealthy landlord is more at ease taking sickness leave to recover from health risks, because his or her income does not rely on labor-intensive production as much as a landless individual. In fact, the landlord is more capable of coping with all labor-related risks such as involuntary unemployment, fluctuations in wage rates, and seasonality in the labor market than the landless because of the relative independence from labor-intensive activities of his or her income sources. In addition, since land can be used as collateral in informal borrowings, it helps reduce the interest rate of the loan or increase the amount of loan taken out by reducing default risks. Similarly, when insurance is imperfect, incomes and family demographic characteristics also affect the scope and coverage of mutual insurance an individual receives, because they determine the ability of making repayments. For instance, an individual from a family with more working adults is more capable of engaging in a mutual insurance arrangement because the contingent income transfers is backed up by more income-earners.

There are two concerns associated with including wealth in X and Y as an explanatory variable. Wealth is accumulated over time and endogenous, and, the marginal utilities of consumption and health may be higher for wealthier individuals, which may still be compatible with perfect risk-sharing. The implication of the first concern is that the significance of wealth in the consumption or sickness leave equation cannot be interpreted as evidence against full insurance. In the empirical estimations, land is used as a proxy for wealth, because it represents at least 80 percent of family assets. The ICRISAT data set provides a categorical variable that classifies landholding into four classes—landless, small farm, medium farm, and large farm—and the landholding classes never changed during the survey years. Therefore, I treat landholding as a household fixed factor in panel regressions.

The correction for the second concern is also clear. When there is no restriction on the functional form of the utility function, we can at most say that the weighted (by welfare weights) marginal utilities of consumption, health, and leisure for all individuals are equalized in any given state, and that the weighted (by the probability) marginal utilities are equalized for each individual across all states. In principle, the marginal utilities can be expressed completely in terms of all exogenous factors in the economy under study. To proceed with empirical tests for full insurance, the only approach is reduced-form estimation, that is, regressing consumption levels on a set of explanatory variables, exogenous and endogenous, and using instrumental variables to control for endogeneity. In the case of wealth, since it is endogenous, there is an endogeneity problem if it is included directly on the right-hand side when estimating individual consumption. In a semi-structural system such as (6.3), the endogeneity of wealth is not a problem. Under full insurance, other things being equal, the consumption of an individual is the same as the average consumption of the village. Other endogenous variables do not matter, since they are assumed to be captured in the village average consumption. If insurance is imperfect, the difference between the consumption of an individual and the average consumption of the village is no longer independent of other endogenous variables, including wealth. As discussed in Chapter 3 and Chapter 5 of this dissertation, wealth restricts the scope and coverage of insurance that an individual receives, therefore affects his or her consumption across various states of nature. In (6.3) and later empirical formulations, the dependent variable is not individual consumption, but rather, the difference between individual consumption and village average consumption. In other words, we have already controlled for wealth effects on the marginal utility of consumption of an individual by using the envelope theorem. The "residual effects" that wealth has on this difference shall reflect imperfect insurance.

The set of equations in (6.3) is a semi-structural form for the equilibrium conditions in Chapter 3 of this dissertation. All of the equilibrium conditions are in terms of the actual levels of consumption, sickness leave, and leisure, not about marginal utilities. This is the benefit of being explicit about the utility specifications—actual levels of the utility commodities can be solved in terms of village aggregates and individual-specific factors.

The error terms u_{it} and v_{it} are functions of the idiosyncratic health risks according to the theories developed in Chapter 3 of this dissertation. However, in an empirical formulation such as (6.3), they not only represent the unobserved idiosyncratic health risks, but also include other factors, characteristics, measurement errors, and risks either not included in theoretical discussions (omitted variables due to mis-specification) or unobservable to the econometrician.

A drawback here is that all errors and risks are summarized under u_{it} and v_{it}, so that it is empirically impossible to identify the actual sources that cause consumption and health fluctuations. For example, sickness leave fluctuations may result from weather changes or unexpected changes in market price of agricultural products, not health risks. One way of separating the aggregate shocks from the idiosyncratic shocks is to include a village-year dummy in the right-hand-side of (6.3) for each panel regression. Notice that the village average consumption and sickness leave will also be included in the village-year dummy, in addition to the aggregate shocks. (6.3) can be rewritten as

$$(6.3)' \quad \begin{cases} C_{it} = \alpha_i^1 + \beta^1 D_t^1 + \gamma_i^1 T_{it} + \delta_i^1 A_{it} + \varsigma_i^1 X_{it} + u_{it} \\ T_{it} = \alpha_i^2 + \beta^2 D_t^2 + \gamma_i^2 C_{it} + \delta_i^2 A_{it} + \varsigma_i^2 Y_{it} + \eta_i w_{it} + v_{it} \end{cases},$$

where the D's are the village-year dummies.

Not knowing the actual nature of the study individuals' preference, there are two sets of null hypotheses that we can consider in (6.3) and (6.3)'. One null hypothesis is $\gamma = 0$ and $\varsigma = 0$ jointly, that is, preferences are separable and insurance is perfect. The other null hypothesis, $\varsigma = 0$, is less restrictive on preferences. The implications of testing full insurance under general nonseparable preferences are discussed in greater detail in the next section.

6.2.2 EXCESS SENSITIVITY IN CONSUMPTION AND SICKNESS LEAVE

Recall that in Chapter 3 and Technical Appendix B, we find that under general nonseparable preferences, the relative percentage change in health and consumption depends on $\frac{\xi_{HC}}{\xi_{CH}}$, other things equal. $\frac{\xi_{HC}}{\xi_{CH}}$ describes the curvature of the utility function, or, the risk attitude of an individual. Conditioning on other factors, if consumption appears to vary more than this ratio in the determination of health, we can reject full insurance and infer how much welfare loss is incurred due to imperfect insurance. Similar arguments apply to the excessive variations in sickness leave in the determination of consumption.

The tests for full insurance can then be carried out in panel regressions, in which individuals from the same village are pooled together. Let us choose the point of log linear approximation as the village average. Individual *i*'s consumption variation from the village average is given by $\Delta C_{it} = C_{it} - \overline{C}_t$. Of this total variation, $\frac{\xi_{HC}}{\xi_{CH}}$ proportion is due to nonseparability in consumption and health, which is an individual-specific factor (see Technical Appendix B). The rest, if there is any, is due to imperfect insurance, therefore is a function of the socioeconomic status of the individual, because these factors restrict the scope and coverage of mutual insurance the individual receives if insurance is imperfect. Similarly, we can derive the variation of individual *i*'s sickness leave from the village average, $\Delta T_{it} = T_{it} - \overline{T}_t$. According to the derivations in Chapter 3 and Technical Appendix B of this dissertation, $\frac{\xi_{HC}}{\xi_{CH}\sigma}$ proportion is due to nonseparability in consumption and health, which is an individual-specific factor. The rest of the variations in sickness leave are due to imperfect insurance, which is explained by the socioeconomic status of the individual.

To sum up, tests for full insurance can be conducted by the following semi-structural panel regression:

$$(6.4)\quad \begin{cases} \left(1-\frac{\xi_{HC}^{i}}{\xi_{CH}^{i}}\right)\left(C_{it}-\overline{C}_t\right)=\alpha_i^1+\varsigma_i^1 X_{it}+u_{it} \\ \left(1-\frac{\xi_{HC}^{i}}{\xi_{CH}^{i}}\right)\left(T_{it}-\overline{T}_t\right)=\alpha_i^2+\varsigma_i^2 Y_{it}+\eta_i\left(w_{it}-\overline{w}_t\right)+v_{it} \end{cases}$$

when aggregate shocks and idiosyncratic shocks are pooled together in u_{it} and v_{it}, and

$$(6.4)'\quad \begin{cases} \left(1-\frac{\xi_{HC}^{i}}{\xi_{CH}^{i}}\right)C_{it}=\alpha_i^1+\beta^1 D_t^1+\varsigma_i^1 X_{it}+u_{it} \\ \left(1-\frac{\xi_{HC}^{i}}{\xi_{CH}^{i}}\right)T_{it}=\alpha_i^2+\beta^2 D_t^2+\varsigma_i^2 Y_{it}+\eta_i\left(w_{it}-\overline{w}_t\right)+v_{it} \end{cases}.$$

Testing full insurance is equivalent to testing $\varsigma = 0$ for the socioeconomic factors X and Y. The coefficient on the mean wage difference, η_i, is expected to be positive under the theoretical predictions from Chapter 3.

An empirical difficulty with (6.4) or (6.4)' is that we do not observe $\frac{\xi_{HC}}{\xi_{CH}}$ and $\frac{\xi_{HC}}{\xi_{CH}\sigma}$ from the data. Let us consider a reduced form of (6.4):

$$(6.5)\quad \begin{cases} C_{it}-\overline{C}_t=\alpha_i^1+\beta^1\overline{C}_t+\gamma_i^1 T_{it}+\delta_i^1 A_{it}+\varsigma_i^1 X_{it}+u_{it} \\ T_{it}-\overline{T}_t=\alpha_i^2+\beta^2\overline{T}_t+\gamma_i^2 C_{it}+\delta_i^2 A_{it}+\varsigma_i^2 Y_{it}+\eta_i(w_{it}-\overline{w}_t)+v_{it} \end{cases}$$

(See Technical Appendix B for derivations).

Similarly, the reduced form of (6.4)' is given by

$$(6.5)'\quad \begin{cases} C_{it}-\overline{C}_t=\alpha_i^1+\beta^1 D_t^1+\gamma_i^1 T_{it}+\delta_i^1 A_{it}+\varsigma_i^1 X_{it}+u_{it} \\ T_{it}-\overline{T}_t=\alpha_i^2+\beta^2 D_t^2+\gamma_i^2 C_{it}+\delta_i^2 A_{it}+\varsigma_i^2 Y_{it}+\eta_i(w_{it}-\overline{w}_t)+v_{it} \end{cases}.$$

The sickness leave of an individual is included in the right-hand side of the consumption equation, because $\frac{\xi_{HC}}{\xi_{CH}}$ varies with health. Other things equal, the coefficient γ_i^1 only reflects the degree of substitutability between consumption and health, therefore shall not be used as a test for full insurance. Similarly, the consumption level of an individual is included in the right-hand side of the sickness leave equation, because $\frac{\xi_{HC}}{\xi_{CH}\sigma}$ varies with consumption. Controlling for other factors, γ_i^2 measures the substitutability between consumption and health and the net rate of investment in health. Therefore, it shall not be used as a test for full insurance either. The real test for full insurance is still $\varsigma=0$.

6.3 EMPIRICAL RESULTS

Full insurance can be tested in reduced forms by (6.3) or (6.3)'. In fact, (6.3) or (6.3)' can serve purposes more than testing full insurance. For example, it can be used to test whether or not the income effects dominate the substitution effects by testing the statistical significance of the coefficient on wage rate, η_i. It can also be used to test the separability between consumption and health by testing the joint significance of the γ's and the ς. If $\gamma = 0$ and $\varsigma = 0$ is rejected jointly, we can reject the null hypothesis that preferences are separable and insurance is perfect. However, since there are more than one alternative hypothesis against this null, further tests are needed in order that we are able to infer more on preferences and insurance. For example, if $\gamma = 0$ and $\varsigma = 0$ are rejected jointly but $\varsigma = 0$ is not rejected in a separate test, then full insurance is not rejected, and the coefficient γ can be used to infer the degree of nonseparability between consumption and health in preferences. If $\gamma = 0$ and $\varsigma = 0$ are rejected as well as $\varsigma = 0$ in a separate test, one can reject full insurance, however, the coefficient γ may reflect the tradeoff between consumption and health due to imperfect insurance (the excess responses), and the changes between consumption and health due to nonseparability. To separate these two effects, (6.5) or (6.5)' can be used.

In the following, consumption and sickness leave from work are estimated jointly in panel regressions to test for full insurance with the ICRISAT data. All consumption and grain consumption are estimated separately, because grain consumption is the key to food security and nutrition. Information on all consumption items was collected at the household level. To circumvent the data aggregation problem, adult-equivalent consumption measures are used in regressions. They are the age and sex adjusted consumption of an adult-equivalent in a household (see Chapter 2, 3, and 5 of this dissertation for more discussions).

As discussed in Chapter 5 of this dissertation, even though one is able to use the complete ten-year data set, there are selection problems with the labor data. To obtain consistent estimates, the results from Chapter 4 of this dissertation are incorporated here. The results from Chapter 4 suggest that conditioning on household demographic characteristics, an individual from a lower income household is more active in the village labor market. Therefore, when estimating consumption and sickness leave jointly to test full insurance, we shall also include a third equation for income to control for endogeneity problems. Here, the adult-equivalent income of a household is regressed on the previous year's average adult-equivalent food consumption of the village.

Consumption and income measures are deflated by a set of village-specific prices developed by Walker and Ryan (1990). (6.3), (6.3)', (6.5), and (6.5)' are estimated with 3SLS. The results are presented in Tables 6.1.a, 6.1.b, 6.2.a, and 6.2.b respectively.

6.3.1 TESTING FULL INSURANCE

Explanatory Variables

There are three categories of explanatory variables included in the regressions, namely, measures of household and individual demographic characteristics, socioeconomic factors that may restrict the scope and coverage of mutual insurance, and variables that appear in the structural equations in (6.3) or (6.3)'. The right-hand-side explanatory variables included in the consumption equation are not all the same as those included in the sickness leave equation.

The age and sex of an individual, and the interaction terms controlling for family structure are not included in the consumption equation because the consumption measures are in adult-equivalent terms, which can be viewed as the age-sex adjusted average consumption of a household. They are included in the sickness leave equation because sickness leave is measured at individual level. They also serve as a check for intrahousehold allocation of labor resources and can be used to infer how risk-sharing is achieved—how much is achieved within a household, and how much is resorted at the village level. Wage rate is included in the sickness leave equation, because the theories developed in Chapter 3 of this dissertation predict that it negatively affects one's sickness leave. It is not included in the consumption equation because consumption is measured annually, so that all impacts from wage rate changes have already been summarized in adult-equivalent income on the right-hand side.

In the consumption equation, the number of family size, the number of age 18 to 45 adults, and the number of age 17 or less children are included as measures of household demographic characteristics. Individual level demographic characteristics such as age and sex are not included, because the dependent variable is the adult-equivalent consumption of the household, which is already adjusted by age and sex. The proportion of yearly sickness leave is included as a measure of nonseparability since it appears in the structural equation. It can also be used to infer the tradeoff between consumption and health if there is imperfect insurance. The adult-equivalent income of the household and the proportion of involuntary unemployment are included as measures of the socioeconomic status the household. Recall in Chapter 5 of this dissertation that the proportion of involuntary unemployment in univariate panel regressions in Townsend's fashion is found to be positive and significant in the consumption equations. It may either be attributed to the nutrition-based efficiency wages, or result from income transfers from other households to offset the adverse impacts on consumption due to unemployment, or both. With multivariate panel regressions, we now are able to control for other factors and separate these two effects.

In the sickness leave equation, the dependent variable is the ratio of reported incidences of sickness leave out of total interview times for an individual. Age and sex of an individual and the aforementioned household factors are included as measures of demographic characteristics. Interaction terms such as the interaction between the age and sex of the individual under study and the number of adults in the household, and the interaction between the age and sex of the individual under study and the number of children in the household are also included. They can serve as checks and controls for intrahousehold allocation of economic resources. Factors such as landholding classes and the household adult-equivalent income are included because they may affect the scope and coverage of mutual insurance a household receives. Wage rate appears in the structural equation for sickness leave, but it is not included in the regressions. The village average wage rate is used instead, because it is the average opportunity cost an individual incurs when taking sickness leave. The adult-equivalent consumption of a household is included in the sickness leave equation, serving as a measure of nonseparability between consumption and health.

A selection factor is also included in the sickness equation. It is the ratio between the predicted interview frequencies and the actual interview frequencies. From Chapter 4 of this dissertation, we know that the number of interviews an individual receives depends on three factors—individual demographic characteristics determining the type of job and nature of employment, family structure and socioeconomic status determining the intrahousehold allocation of labor resources, and the learning-by-doing of the investigator. The predicted value of interview frequencies summarizes all of these factors. Its relative value to the actual interview frequencies summarizes how these factors make the labor market observability of an individual deviate from the village average.

Consumption Determination

In the 3SLS regressions for mean-differences of consumption (Table 6.1.a), we can see that the coefficients on household adult-equivalent income are positive and significant for all three villages, with an average of 0.265, and the magnitudes are very close to each other. In the 3SLS regressions for consumption with village-year dummy variables (Table 6.1.b), we find the estimates are of similar magnitudes, with an average of 0.236. This is strong evidence against full insurance under all preference specifications. Other things equal, the results suggest that around 25 percent of consumption is due to income fluctuations. Following the calculations in Chapter 1 of this dissertation, with a moderate relative risk aversion of 2, we can infer that households are willing to pay around 6.6 percent of their income to eliminate income fluctuations and protect consumption, an amount approximately equals to a medium farm household's annual profits from animal husbandry.

Sickness leave is positive and significant for all three villages in the mean-difference as well as village dummy regressions. The estimates are the highest for Aurepalle and the lowest for Kanzara in both regressions. As discussed earlier in this chapter, these cannot be taken readily as evidence against full insurance, because possible nonseparability between consumption and health has not been controlled for.

A test for joint significance of the coefficients on adult-equivalent income and sickness leave in the mean-difference consumption equation is performed for each village. The $\chi^2(2)$ statistics for Aurepalle, Shirapur, and Kanzara are 954.92, 3905.81, 3010.79, respectively, rejecting the hypothesis that the coefficients are jointly insignificant at 95% confidence level. Given that full insurance is reject on the grounds that the coefficients on adult-equivalent income for all villages are positive and significant, we can reject the hypothesis that preferences are separable and insurance is perfect. To infer how much tradeoff between consumption and health is incurred due to imperfect insurance, we need to use a specification such as (6.5) or (6.5)'.

Similar patterns can be found in the regressions involving village-year dummies, although with much larger magnitudes. A test for joint significance of the coefficients on adult-equivalent income and sickness leave is performed for each village, and the $\chi^2(2)$ statistics for Aurepalle, Shirapur, and Kanzara are 3311.59, 4385.00, 1876.33, respectively, rejecting the hypothesis that the coefficients are jointly insignificant at 95% confidence level.

In the mean-difference regressions, household demographic structure is found to affect the adult-equivalent consumption of a household in the villages of Shirapur and Kanzara, but not in Aurepalle. Recall that the adult-equivalent consumption is the weighted-average per capita consumption of a household. Under full insurance, an average individual receives the average of village consumption, other things equal. The estimates on family size for Shirapur and Kanzara clearly do not support this view. As family size increases, an average individual receives lower consumption because of imperfect insurance, suggesting that the village mutual insurance has limited coverage, especially for those households who have more family members to support. For example, the annual reduction in adult-equivalent consumption with the addition of a family member is around 55.2 rupees for the households in Shirapur, and around 66.3 rupees for the households in Kanzara, other things equal.

The number of age 18 to 45 adults seems to have a positive impact on the adult-equivalent consumption of a household, also in contradiction to the full insurance predictions. It implies that with limited coverage from the village mutual insurance, households with more working adults are more capable of coping with idiosyncratic risks and smoothing consumption. For example, in Shirapur, the results suggest that the weighted per capita consumption of a household increases by 53.6 rupees annually if the household adds one addition working adult. In Kanzara, this number is around 27 rupees.

The number of age 17 or less children is found to be positive and significant on the adult-equivalent consumption of a household, suggesting that other things equal, as the number of dependents increase, the consumption allocated to an adult-equivalent also increases, other things equal. However, this cannot be taken as evidence for or against insurance, because children are usually given priorities in consumption. To sum up, the significance of household demographic characteristics provides evidence against full insurance at the village level, as well as evidence for intrahousehold allocation of economic resources. Similar results can be found in the regressions with village-year dummies, and the pattern is more consistent and the magnitudes are also much larger.

One of the most important critiques on the empirical tests for complete risk sharing regards the treatment of aggregate shocks. It argues that part of the comovement among individual choice variables is driven by aggregate shocks. Therefore, if one does not control for aggregate shocks, the degree of risk sharing may be over-estimated. The differences in Table 6.1.a and 6.1.b clearly is in support of this critique. The difference between (6.3) and (6.3)' (i.e., the difference between Table 6.1.a and 6.1.b) is that (6.3)' controls for all village-level factors, including aggregate economic resources, rainfall, price changes in the village markets, infrastructure, environmental health problems, etc., while (6.3) only controls for village-level aggregate economic resources. For example, in the regressions with village-year dummies where all village-level factors are controlled for, the adult-equivalent consumption is found to reduce by 65 rupees with the addition of a family member for a household in Shirapur, and to increase by 70 rupees if the household has one more working adult. In the mean-difference regressions, the corresponding estimates are 55 rupees, and 53 rupees. Conditioning on other factors, we can infer that around 10 percent of the variations in individual consumption is due to changes in village-level conditions rather than aggregate economic resources.

Similar to the findings in Chapter 5 of this dissertation, the coefficients on the proportion of involuntary unemployment tend to be positive and significant in the regressions, at least for Shirapur and Kanzara. The estimates in the regressions with village-year dummies share similar features although much smaller. The reason for this difference is similar to the above—part of the variations in the proportion of involuntary unemployment is due to village-level factors such as annual rainfall, technological changes, and government projects. These are included in (6.3) together with idiosyncratic variations, which results in larger estimates in Table 6.1.a.

Since the impacts of income, sickness leave, and household demographic factors have been controlled for in the regressions, the coefficients on involuntary unemployment can be viewed as reflecting mutual insurance. Other things equal, we can infer that consumption is protected at the incidence of involuntary unemployment through the informal mutual insurance arrangements.

Sickness Leave

The coefficients on the adult-equivalent household income are found to be negative and significant in the mean-difference regressions for all villages, however, they are not as significant as in the regressions with village-year dummies. For example, the estimate for Shirapur is –0.024 in the mean-difference regressions, and –0.012 in the regressions with village-year dummies. Controlling for other factors, they suggest that with 100 rupees increase in adult-equivalent household income, if there is no changes in the overall economic conditions, the average sickness leave decreases by 2.4 percent in Shirapur. Controlling for fluctuations in overall economic conditions in Shirapur, this percentage is about half as before.

According to the theories developed in Chapter 3 of this dissertation, consumption and health are both normal goods, therefore, when income increases, an individual is expected to consume more and take longer sickness leave, other things equal. The estimates on adult-equivalent income from both regressions seem to be in contradiction to the theoretical predictions. It might be that there exist other linkages between income and health that are not considered in the theories. For example, higher incomes can lead to better sanitation and hygiene, and improve preventive cares, therefore reduce the incidence of human illnesses. In theoretical terms, income changes the distribution of the idiosyncratic health risks. Conditional on other factors, an individual with a higher income has either lower expected health risks, or lower variations in health risks, or both. Village-level risk sharing is not longer based on a single distribution of health risks. It becomes a multi-leveled process. If this is the case, the significance of adult-equivalent income shall not be taken as evidence for or against full insurance, because we do not know the actual multi-leveled distributions of health risks.

The coefficients on landholding class in the mean-difference regressions (large farm is the omitted category) for Aurepalle and Kanzara are negative and significant, suggesting that large farm households take much longer sickness leave relative to others in the same village, other things equal. In the regressions with village-year dummies, however, the estimates are very noisy. In the mean-difference regressions, the results suggest that conditioning on other factors (including fluctuations in overall economic conditions), as landholding class increases, the difference in the proportion of yearly sickness leave decreases relative to the large farm households. Take, for example, in Aurepalle, individuals from landless households take 29.5 percent shorter sickness leave than those from large households in the same village, other things equal. This difference becomes 23.2 percent for those from small farm households, and is 14 percent for those from medium farm households.

Landholding status determines the income sources of a household. Landless households derive at least 80 percent of incomes from the village labor markets. As landholding increases, a household relies less on labor incomes. When afflicted with illnesses and injuries, those who are more dependent on labor incomes find it more expensive to take sickness leave when insurance is imperfect, other things equal. Therefore, we can reject full insurance because the decision to take sickness leave is affected by a household's socioeconomic status. The results also suggest that conditioning on other factors, when afflicted with similar health problems, a landless farmer has a lower health status than a large farmer, because he or she seek less sickness leave (the metaphor for medical treatment).

Family size is found to be negative and significant for Aurepalle and Kanzara in the mean-difference regressions, but its direction and significance are not very clear in the regressions with village-year dummies. Nonetheless, we can reject full insurance in Aurepalle and Kanzara. Under full insurance, the amount of sickness leave solely depends on the severity of the health problems. However, here we see that household structure affects this decision. Conditioning on other factors, it suggests that with an additional family member, an individual reduces the amount of sickness leave by 3.3 percent in Aurepalle, and 22.5 percent in Kanzara.

To learn more about the nature of intrahousehold allocation of labor activities, two additional explanatory variables are included in the sickness leave equation—the interaction term between the age and sex of the individual under study and the total number of working adults in the family, and the interaction term between the age and sex of the individual under study and the total number of children in the family. The basic idea is to check how a bread-earner's decision of taking sickness is affected by the number of dependents and the number of total bread-earners in the family. When the village mutual insurance has limited scopes and coverage, households have to rely more on their own, therefore, we expect to see that factors determining intrahousehold allocation be play significant roles in determining sickness leave. The empirical findings here seem to support this view. The results from the mean-difference regressions suggest that with one more working adult in the family, a working adult takes longer sickness leave by 0.65 percent in Shirapur, and by 5.64 percent in Kanzara, other things equal. The results from the mean-difference regressions suggest that with one more child to support in the family, a working adult takes shorter sickness leave by 2.49 percent in Shirapur. Similar patterns can be found in the regressions with village-year dummies, however, the magnitudes are much smaller, suggesting that at least half of the intrahousehold allocations of labor activities are responses toward changes in overall economic conditions.

Village average wage rate is included in the sickness leave equation in the mean-difference regressions because wage rate appears in the structural equation for the determination of sickness leave (see Chapter 3 of this dissertation). It is dropped in the regressions with village-year dummies, because it is unidentifiable from the dummies. Village average wage rate is used here instead of individual wage rate because it represents the average opportunity cost if one takes time off from work to seek treatment. It is also because the village labor market is operated on a daily basis, so that individuals can engages in different types of jobs across different days within a short time period. According to the theoretical prediction, as wage rate increases, an individual takes shorter sickness leave. However, the empirical findings on wage rate are not very clear. In the mean-difference regressions, the coefficient on wage rate is negative and significant for

Shirapur, but positive and significant for Kanzara. This may result from measurement errors in wage rate.

The coefficients on adult-equivalent consumption are positive and significant in the sickness leave equation in the mean-difference regressions, although very noisy in the regressions with village-year dummies. Since we reject full insurance on the grounds that the socioeconomic status and demographic characteristics of a household affect an individual's decision to take sickness leave, we can infer that at least part of the sickness leave response is due to imperfect. However, to see how much tradeoff between consumption and health is incurred due to imperfect insurance, we need to explore further with specification such as (6.5).

Selection factors are included in the sickness leave equation for both mean-difference and village-year dummy regressions. However, none of the estimates are significant. The main reason is that the selection factor itself summarizes household demographic characteristics, incomes, and the learning-by-doing of the investigator, which have already been included in the 3SLS estimations—household demographic characteristics are included directly in the sickness leave equation, incomes are regressed on past year's adult-equivalent food consumption separately to control for endogeneity, and the learning-by-doing effects are account for the village-year dummies.

6.3.2 NONSEPARABILITY AND EXCESS SENSITIVITY

This section re-estimates consumption and sickness leave jointly using 3SLS with (6.5) and (6.5)'. The results are presented in Tables 6.2.a and 6.2.b. The focus is to detect excess variations in consumption and sickness leave due to imperfect insurance under a general nonseparable preference specification. (6.5)' accounts for overall economic fluctuations, therefore, it sheds more light on the effectiveness of the informal risk sharing among the villagers.

If preferences are separable in health and consumption, the coefficient on the village average consumption, $\overline{C}_t$, is equal to zero in (6.5) in the consumption equation, and the coefficient on the village average sickness leave, $\overline{T}_t$, is also zero in (6.5) in the sickness leave equation. The coefficient on sickness leave in the consumption equation in (6.5) is equal to that in (6.3), and so is the coefficient on consumption in the sickness leave equation. The coefficients on other explanatory variables are expected to be similar to those obtained in (6.3). Similar arguments can be made to (6.3)' and (6.5)'.

Under the alternative that preferences are not separable in health and consumption, the coefficient on $\overline{C}_t$ in the consumption equation and the coefficient on $\overline{T}_t$ in the sickness leave equation are equal to $1-\frac{\xi^i_{HC}}{\xi^i_{CH}}$ and $1-\frac{\xi^i_{HC}}{\xi^i_{CH}\sigma}$ respectively. They reflect the excess responses in consumption and sickness leave due to variations in the socioeconomic status of a household.

In the mean-difference regressions, the coefficients on $\overline{C}_t$ are negative and significant except for Kanzara, showing a strong tendency of rejecting the null hypothesis that preferences are independent over consumption and health. In the regressions with village dummies, the counterpart to $\overline{C}_t$ is the village-year dummy D_t. The estimates on D_t for Shirapur and Kanzara are all negative and significant relative to the omitted year of 1975. The estimates on D_t for Aurepalle are all positive and significant relative to the omitted year of 1975. Note that the estimates on D_t are not comparable with the estimates on $\overline{C}_t$, since they only tell the difference of a particular year from a base year. Regardless of the signs, as long as the estimates are significant, we can reject the null that consumption and health are independent over consumption and health.

The coefficients on $\overline{T}_t$ show a similar tendency in the mean-difference regressions, although not as clear. $\overline{T}_t$ is suppressed by the village-year dummies, therefore not reported in Table 6.2.b. They seem to suggest that marginal rate of substitution between consumption and health, $\frac{CU_C}{HU_H}$, is greater than 1 around the village average when insurance is imperfect (see Technical Appendix B). In other words, health has a lower value relative to consumption for the study households, and improved health status by taking more sickness leave can increase welfare. Notice that the theories and empirical estimations are all based on a static model. In reality, however, health functions as a true stock variable in the sense that current low health status lowers future labor productivity. Therefore, persistently low health status due to imperfect insurance incurs much larger welfare losses in a dynamic setting.

Estimates on household adult-equivalent income, involuntary unemployment, wage rate, the number of children in the family, and family size are all very similar to those obtained in (6.3) and (6.3)'. The estimates on landholding classes are not very consistent with the previous results, possibly due to the inclusion of the additional explanatory variables.

As before, selection factors do not have any significance in the regressions since the key factors determining the selection biases are already included in the regressions.

6.4 CONCLUSIONS AND POLICY IMPLICATIONS

The empirical results derived in this chapter seem to fit the theoretical implications very well. Tests for full insurance under separable preferences are similar to those under nonseparable preferences, because consumption and sickness leave have similar properties under these two regimes. This is why reduced forms such as (6.3) can be used to as a unified test for full insurance in a general framework where there is no restriction on the substitutability between consumption and health. (6.5) is a further development based on (6.3). It not only carries on the role of (6.3) as tests for full insurance, it also can demonstrate the degree of nonseparability between consumption and health, and the tradeoff between consumption and health due to imperfect insurance, namely, the excess responses.

One of the common critiques on the tests for risk sharing is that comovement of choice variables across individuals or households are partly driven by aggregate shocks. (6.3)' and (6.5)' are further developments based on (6.3) and (6.5) because they answer this critique by including a set of village-year dummy variables to account for overall economic fluctuations.

In addition to imperfect insurance and nonseparability between consumption and health in preferences, a very important empirical finding in this chapter is that the study households do not take enough sickness leave when afflicted with health problems, and that taking longer sickness leave can help improve welfare. In other words, the study individuals have under-invested in health. Even when the mutual insurance arrangements are not perfect, it is still possible to improve welfare by increasing the amount of sickness leave relative to consumption. What might have caused the under-investment in health? One explanation is that the individuals are myopic—the discounted value of future health is unreasonably low. Another explanation is that individuals are not aware of the importance of good health today and good health in the future. Having realized these, public policies aiming at poverty reduction, public health, and efficiency of financial services can be carried out more efficiently.

In *Investing in Health* (World Development Report, 1993), the World Bank introduced the concept of Disability Adjusted Life Years (DALYs) and estimated that the global disease burden was unevenly spread and concentrated in low income and middle income countries. In particular, sub-Saharan Africa, China, and India account for 60 percent of the global disease burden. The report identifies four major problems with international healthcare systems: the misallocation of funds to less cost-effective interventions, the ineffective use of funds, the inequality in access to basic healthcare, and the explosion of healthcare costs outpacing the growth in income. The Bank recommended that low income and middle income countries should shift the focus of government investment away from tertiary healthcare toward public health, promote diversity and competition by introducing private or social insurance schemes, and foster competition in the delivery of health services, and improve family health by educating and empowering women. The bank argued that a healthy economy needed a healthy workforce, and that, of possible economic policies, "... increasing the income of those in poverty is the most efficacious for improving health. The reason that the poor are most likely to spend additional income in ways that enhance their health: improving their diet, obtaining safe water, and upgrading sanitation and housing."

The success of a health-improving policy may not be easily seen until after some fifteen or twenty years, which seems to suggest that developing countries should carry out public health policies as soon as possible in order to benefit from them in the future. However, before carrying out public health policies, we shall at least understand the nature of market failures and the efficiency of the existing private institutions. In our case, the study households face tremendous risks in production, income and health, and are found to engage in informal risk-coping arrangements when formal institutions do not exist. Clearly the informal arrangements do not provide perfect insurance to protect consumption and health—we find that socioeconomic factors such as income, wealth, and family structure restrict the scope and coverage of risk protections, also, due to myopic and unawareness, health is seriously under-invested. For the ICRISAT households, policy interventions shall be carried out in two dimensions—targeting the less well insured by offering additional credits and insurance, and reckoning the awareness of the importance of health through education, either educating the household head since he or she is the key decision-maker of a family, or educating the mother so that she is able to care for her children in a better way, or educating the general public about environmental health, transmissions of diseases, and prevention can also reduce the occurrence of covariate health risks.

Health and consumption are found to be nonseparable in preferences for the study households. One possible explanation is that they translate into each other through the nutrition link. It seems to be consistent with the findings of the existing literature on health and nutrition, that at least for very low income individuals, better nutrient- and caloric-intake improve one's health. If this indeed is the case, policies of food and nutrient subsidies will also help improve welfare greatly.

Table 6.1.a 3SLS with Mean-Difference (Full Insurance)

Mean-Difference of Consumption and Sickness Leave						
Village	Aurepalle		Shirapur		Kanzara	
Variable	Consump -tion Eq.	Sick Leave Eq.	Consump-tion Eq.	Sick Leave Eq.	Consump-tion Eq.	Sick Leave Eq.
Adult-equivalent Consumption		0.116* (0.006)		0.028* (0.006)		0.210* (0.039)
Proportion of Year Sick Leave	18.303* (0.637)		8.8645* (0.462)		1.635* (0.450)	
Adult-equivalent Income	0.239* (0.003)	-0.024* (0.001)	0.294* (0.004)	-0.024* (0.002)	0.291* (0.005)	-0.149* (0.028)
Landholding						
Landless		-29.476* (2.383)		-0.405 (3.810)		-20.918* (7.854)
Small Farm		-23.209* (1.895)		-1.053 (1.241)		-17.739* (6.820)
Medium Farm		-13.989* (1.238)		-0.130 (3.084)		-15.751* (5.874)
Age		-0.078* (0.014)		0.126* (0.022)		0.800* (0.236)
Sex =1 if male		5.528* (0.603)		-0.707 (1.013)		9.011* (5.621)
# Members	31.104 (19.512)	-3.340* (0.873)	-55.197* (9.8129)	1.546 (0.950)	-66.324* (14.936)	-22.470† (12.404)
# Adults	-30.447 (19.517)	1.900* (0.859)	53.633* (10.282)	-1.861† (1.010)	27.135† (15.954)	20.202† (10.575)
# Children	-6.938 (18.842)	-0.861 (0.836)	65.421* (10.510)	-2.490* (1.245)	93.901* (15.929)	3.0797 (6.724)
Adult × # Adults		-0.091 (0.190)		0.653† (0.342)		5.642* (2.842)
Adult × # Children		-0.252* (0.064)		-0.0840 (0.0789)		-1.5609* (0.7940)
Proportion of Involuntary Unemployment	-124.137 (79.383)		373.580* (66.993)		732.830* (133.682)	
Wage Rate		34.754 (24.812)		-9.247* (4.580)		140.912* (53.084)
Selection Factor		-0.107 (0.446)		0.018 (0.552)		-0.125 (0.998)

Note:

1. Standard errors are reported in parentheses. * are significant at 95% confidence interval. † are significant at 90% confidence interval.
2. Adult-equivalent income of a household is instrumented by the household's adult-equivalent food consumption of the previous year.

Table 6.1.b 3SLS with Village Dummy (Full Insurance)

Consumption and Sickness Leave with Village Dummy						
Village	Aurepalle		Shirapur		Kanzara	
Variable	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.
Adult-equivalent Consumption		0.047 (0.032)		-0.037 (0.037)		0.136 (0.174)
Proportion of Year Sick Leave	45.410* (15.300)		17.120* (2.408)		4.345* (1.342)	
Adult-equivalent Income	0.181* (0.009)	-0.009 (0.005)	0.288* (0.006)	-0.012* (0.007)	0.241* (0.005)	-0.044 0.054
Landholding Class						
Landless		-6.705 (8.973)		4.313 (5.506)		-160.715 (104.837)
Small Farm		-5.800 (7.679)		1.709 (0.992)		-151.413 (98.115)
Medium Farm		-4.065 (5.488)		5.003 (4.478)		-111.268 (73.977)
Age		-0.032 (0.041)		0.192* (0.073)		0.422* (0.100)
Sex =1 if male		0.111 (0.277)		-1.361 (0.857)		-2.072 (2.798)
# Members	-81.917* (30.858)	2.521* (1.082)	-65.237* (12.119)	-1.905 (1.995)	-31.946* (16.243)	-24.490† (14.450)
# Adults	79.867* (31.413)	-2.770* (1.444)	70.784* (12.654)	2.038 (2.089)	-0.338 (17.417)	22.200 (16.188)
# Children	71.359* (31.108)	-2.111* (0.902)	53.950* (13.008)	1.961 (2.060)	35.803* (17.577)	8.940* (4.374)
Adult × # Adults		-0.052 (0.158)		0.450† (0.252)		1.853 (5.493)
Adult × # Children		0.026 (0.052)		-0.081 (0.057)		-0.629 (1.416)
Proportion of Involuntary Unemployment	136.454 (167.578)		234.025* (87.772)		668.085* (179.623)	
Selection Factor		0.137 (0.216)		-0.769 (0.691)		3.589 (3.008)

Table 6.1.b Continued

DUM1976	370.031*	-17.270*	106.086*	3.010	176.257*	-8.046
	(53.802)	(8.814)	(26.174)	(5.834)	(35.778)	(21.795)
DUM1977	162.802*	-7.594*	179.613*	2.773	164.781*	3.546
	(55.916)	(2.686)	(29.395)	(8.415)	(34.901)	(11.183)
DUM1978	174.698*	-9.054*	205.579*	2.689	145.419*	-4.805
	(56.312)	(4.191)	(29.571)	(8.758)	(33.937)	(14.086)
DUM1979	246.179*	-10.455*	110.611*	0.293	287.805*	-34.478
	(57.149)	(4.213)	(27.291)	(4.711)	(35.057)	(43.115)
DUM1980	290.742*	-12.155*	140.254*	-0.359	344.472*	-40.511
	(57.764)	(5.548)	(29.297)	(5.475)	(37.528)	(53.343)
DUM1981	258.505*	-9.035*	33.565	-1.519	445.7629	-45.986
	(65.208)	(2.393)	(30.159)	(2.589)	(36.989)	(64.111)
DUM1982	106.450*	-3.165*	-101.298*	-4.618	-43.489	55.027
	(53.386)	(1.504)	(29.518)	(3.193)	(37.467)	(41.081)
DUM1983	-112.202†	8.051	-429.681*	-16.600	-475.859*	150.894
	(64.316)	(7.397)	(30.432)	(16.469)	(38.926)	(138.765)
DUM1984	-128.845*	10.227	-427.677*	-13.550	-313.207*	89.463
	(60.029)	(10.189)	(28.867)	(15.512)	(36.185)	(86.511)

Note:

1. Standard errors are reported in parentheses. * are significant at 95% confidence interval. † are significant at 90% confidence interval.
2. Adult-equivalent income of a household is instrumented by the household's adult-equivalent food consumption of the previous year.
3. 1975 is the omitted year for the village-year dummy variables.

Table 6.2.a 3SLS with Mean-Difference (Excess Sensitivity)

Mean-Difference of Consumption and Sickness Leave						
Village	Aurepalle		Shirapur		Kanzara	
Variable	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.
Village Average Consumption	-1.889* (0.088)		-0.176* (0.024)		0.157* (0.034)	
Average Sick Leave		-0.375 (0.378)		-1.711* (0.535)		-0.593† (0.351)
Adult-equivalent Consumption		0.001 (0.002)		0.006 (0.004)		0.002† (0.001)
Proportion of Year Sick Leave	5.766* (1.129)		12.320* (0.425)		2.086* (0.504)	
Adult-equivalent Income	0.261* (0.003)	-0.0003 (0.0003)	0.328* (0.004)	-0.009* (0.001)	0.328* (0.004)	-0.002* (0.0005)
Landholding						
Landless		1.600 (0.984)		10.757* (1.442)		-2.092 (1.757)
Small Farm		1.721* (0.832)		0.650 (0.774)		-1.482 (1.542)
Medium Farm		0.585 (0.627)		8.778* (1.178)		-0.208 (1.416)
Age		0.025* (0.010)		0.075* (0.014)		0.299* (0.018)
Sex =1 if male		-0.735† (0.404)		-2.596* (0.606)		-1.606* (0.713)
# Members	24.838 (19.593)	0.260 (0.403)	-58.778* (9.906)	1.121 (0.762)	-66.911* (15.089)	1.001 (0.670)
# Adults	-33.371† (19.617)	-0.218 (0.402)	58.307* (10.376)	-1.265 (0.803)	24.098 (16.101)	-1.279† (0.691)
# Children	3.560 (18.944)	-0.303 (0.398)	75.662* (10.581)	-0.758 (0.846)	104.907* (16.129)	-0.345 (0.693)
Adult × # Adults		-0.025 (0.156)		0.614* (0.247)		-0.069 (0.291)
Adult × # Children		0.026 (0.051)		-0.088* (0.055)		-0.090 (0.076)
Proportion of Involuntary Unemployment	399.966* (124.769)		486.355* (50.665)		1181.676* (155.111)	
Wage Rate		3.749 (18.955)		-20.179* (2.246)		4.229 (2.996)
Selection Factor		-0.005 (0.015)		0.224 (0.399)		0.550 (0.025)

Note:

1. Standard errors are reported in parentheses. * are significant at 95% confidence interval. † are significant at 90% confidence interval.
2. Adult-equivalent income of a household is instrumented by the household's dult-equivalent food consumption of the previous year.

Table 6.2.b 3SLS with Village-Dummy (Excess Sensitivity)

Mean-Difference of Consumption and Sickness Leave with Village Dummy						
Village	**Aurepalle**		**Shirapur**		**Kanzara**	
Variable	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.	Consumption Eq.	Sick Leave Eq.
Adult-equivalent Consumption		-0.001 (0.001)		0.077* (0.003)		0.026* (0.003)
Proportion of Year Sick Leave	0.514 (1.210)		11.975* (0.332)		6.559* (0.479)	
Adult-equivalent Income	0.280 (0.004)	0.000 (0.0004)	0.353* (0.004)	-0.027* (0.0009)	0.312* (0.005)	-0.011* (0.001)
Landholding						
Landless		1.324 (1.074)		-2.479* (0.951)		-12.074* (3.360)
Small Farm		1.733† (0.936)		0.366 (0.596)		-11.263* (3.066)
Medium Farm		0.429 (0.706)		-1.782* (0.785)		-7.921* (2.544)
Age		0.024* (0.009)		0.006 (0.012)		0.288* (0.022)
Sex =1 if male		-0.614 (0.393)		0.263 (0.384)		-2.126* (0.722)
# Members	31.310* (19.388)	0.183 (0.407)	-61.593* (10.502)	4.689* (0.811)	-30.573* (15.796)	-0.416 (0.943)
# Adults	-50.106* (19.507)	-0.066 (0.411)	67.845* (10.991)	-5.144* (0.851)	-5.510 (16.839)	0.501 (0.964)
# Children	-23.034 (18.781)	-0.246 (0.400)	53.321* (11.208)	-4.243* (0.871)	41.012* (17.230)	-0.236 (0.881)
Adult × # Adults		-0.029 (0.156)		0.632* (0.307)		0.157 (0.501)
Adult × # Children		0.025 (0.051)		-0.109* (0.068)		-0.145 (0.131)
Proportion of Involuntary Unemployment	197.218* (124.118)		33.208 (38.461)		638.175* (169.671)	
Select		-0.007 (0.406)		-0.356 (0.457)		0.085 (0.849)

Table 6.2.b Continued

DUM1976	889.9713 (43.662)	-0.070 (1.133)	-209.437* (24.237)	-5.709* (1.881)	-330.987* (36.357)	-1.379 (1.739)
DUM1977	878.671* (43.512)	-0.397 (0.945)	-377.420* (29.039)	-6.271* (2.272)	-343.074* (35.839)	0.507 (1.716)
DUM1978	845.690* (44.137)	-0.665 (1.016)	-402.118* (29.524)	-7.046* (2.315)	-208.535* (34.479)	-0.872 (1.653)
DUM1979	767.347* (41.981)	-0.218 (0.949)	-327.453* (26.050)	-1.463 (1.978)	-190.583* (35.928)	-5.773* (1.847)
DUM1980	738.312* (40.599)	-0.271 (0.979)	-399.159* (28.102)	-1.294 (2.137)	-256.280* (38.939)	-6.195* (2.009)
DUM1981	779.282* (41.221)	-0.789 (0.861)	-512.029* (29.398)	5.516* (2.204)	-298.742* (38.850)	-7.190* (2.059)
DUM1982	548.479* (40.410)	-0.516 (0.782)	-517.828* (27.902)	13.971* (2.143)	-477.160* (38.169)	9.369* (2.231)
DUM1983	38.851 (43.833)	-1.299 (0.961)	-618.858* (27.528)	41.439* (2.783)	-678.647* (38.991)	24.501* (3.922)
DUM1984	-230.183* (43.362)	-1.202 (1.151)	-429.659* (25.224)	35.714* (2.512)	-372.997* (36.109)	13.471* (2.706)

Note:

1. Standard errors are reported in parentheses. * are significant at 95% confidence interval. † are significant at 90% confidence interval.
2. Adult-equivalent income of a household is instrumented by the household's adult-equivalent food consumption of the previous year.
3. 1975 is the omitted year for the village-year dummy variables.

CHAPTER 7 CONCLUSION

Coping with risks has been an important part of everyday life ever since human existence. In primitive ages while humans were competing with other species for survival, food and life security were the most imminent. Archaeological and sociological findings reveal that our ancestors managed these risks through storage of food, labor division, and diversification of production activities. In modern times, risks may not be as life-threatening. Individuals with high incomes or living in well-established societies may find risks even less troublesome commercial loans, home mortgages, etc.) and insurance (insurance on life, health, disability, property, unemployment, etc.) that can help them cope with risks very effectively. Individuals with lower incomes or living in less-developed societies are less fortunate because of problems of information flow and contractual enforcement. When formal markets do not function, they have to resort to alternatives.

Most of the existing economic literature on risk-coping strategies has taken the theoretical implications on risk-sharing too mechanically. Perfect risk sharing implies only one thing—welfare smoothing, that is, welfare is smoothed throughout different states. Welfare is a composite of all commodities in the commodity space, including consumption, health, and leisure. Only when consumption is independent from other commodities in the welfare composite in preferences, can consumption smoothing be taken as evidence for Pareto optimum. If consumption is not independent in preferences, testing whether or not consumption is smoothed does not provide much information regarding the effectiveness of the underlying risk-coping strategies, because the so-called "excess sensitivity" in consumption may be compatible with Pareto optimality under nonseparability in preferences. Based on this observation, an important part of this dissertation starts from the most fundamental issues in preferences, and discusses nonseparability and their theoretical implications in the literature of risk sharing. In the theoretical analyses in Chapter 3, it is shown that consumption can be found to have "excess sensitivity" toward changes in health and wage rates under perfect insurance if it is not independent from consumption in preferences. To measure the real excess sensitivity, one needs to peel off the variations in consumption due to nonseparability. Suitable "peeling" techniques are proposed and brought into empirical estimations. Another contribution of the theoretical analysis is that it clarifies the controversies over preference shocks. It reveals that assuming preference shocks is equivalent to assuming some forms of nonseparability in preferences. The shocks are not really shocks to preferences, but rather, changes in the substitution patterns across utility commodities resulting from nonseparability.

Having set the stage for empirical estimations, this dissertation goes on to ask how much protection can the informal mutual insurance provide in protecting individuals against idiosyncratic health risks and overall economic risks. In particular, it asks whether the informal risk-coping strategies can substitute the role of a formal health insurance, so that the individuals are able to seek necessary and sufficient amount of health care when afflicted with illnesses and injuries. It asks who are relatively well protected, and the welfare implications of imperfect insurance.

One strong appeal of the theoretical formulations is that the equilibrium conditions can be put into empirical estimations with little revision. Because of this, empirical estimations are all structural and semi-structural, which allow for interesting findings such as the nonseparability between health and consumption, the degree of excess sensitivity in consumption and sickness leave due to imperfect insurance, and the marginal rate of substitution between consumption and health being greater than 1 which suggests under-investment in health.

This dissertation has also made several important empirical contributions to the literature using the ICRISAT data, including identifying the sample selection problems in the ICRISAT labor data, the revisit of Townsend (1994) with village-specific prices and more years of data.

High quality theoretical and empirical research both requires clear visions, skills, and perseverance. No matter how comprehensive they are, there are still "holes" left over to be mined.This dissertation is by no means an exception. For future research, there are several directions to explore. In Chapter 4, the results show strong evidence for intrahousehold allocation of labor activities, yet all regressions are done in reduced forms without much theoretical discussion regarding the nature of intrahousehold allocation, and its effectiveness in protecting consumption and health relative to the village-level interhousehold resource allocations. How exactly are risks shared? Are they mainly absorbed within a household? How do households choose the types of intrahousehold allocation? In Chapter 3 and Chapter 6, even though preference shocks have been identified as changes in substitution between utility commodities, one can still improve the efficiency of estimations by introducing a random coefficient as a proxy for preference shocks that are nonseparable from the consumption-health composite. The results show that the ICRISAT households have under-invested in health. It is argued that this may result either from myopic behavior or from the unawareness of the importance of health, yet all estimations are in panel. How do the theoretical and empirical results change if one allows for dynamics? What do we mean by risk sharing if risks are carried over time, so are the risk-coping strategies? How to formulate the nutrition link between consumption and health in a dynamic setting? These are important questions that deserve to be explored in further research.

BIBLIOGRAPHY

Baxter, Marianne and Urban Jermann, "Household Production and the Excess Sensitivity of Consumption to Current Income", *American Economic Review*, Vol. 89, No. 4 (1999), 902-920.

Bell, Clive, "Credit Markets and Interlinked Transactions", Chapter 16 in *Handbook of Development Economics*, Vol. 1, edited by Hollis Chenery and T.N. Srinivasan, North-Holland, 1988.

Behrman, Jere and Anil B. Deolalikar, "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India", *Journal of Political Economy*, Vol. 95, No. 3 (1987), 492-507.

Behrman, Jere, "Health and Nutrition", Chapter 14 in *Handbook of Development Economics*, Vol. 1.

_____, "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?," *Oxford Economic Papers*, , Vol. 40, No. 1, (1988), 32-54.

Banerjee, Abhijit V., and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, Vol. 101, No. 2 (1993), 274-98.

Besley, Timothy, "Savings, Credit and Insurance", Chapter 36 in *Handbook of Development Economics*, Vol. 2.

Bidinger, P.D., Nag, B., and Babu, P. "Nutritional and Health Consequences of Seasonal Fluctuations in Household Food Availability", *Food and Nutrition Bulletin*, Vol. 8, No. 1, 36-60.

Cox, Donald and Emmanuel Jimenez, "Social Security and Private Transfers in Developing Countries: the Case of Peru." *World Bank Economic Review*, 1992, 6:1

Dasgupta, Partha and Debraj Ray, "Inequality as a Determinant of Malnutrition and Unemployment: Theory", *Economic Journal*, 96

______, "Inequality as a Determinant of Malnutrition and Unemployment: Policy", *Economic Journal*, 97

Deaton, Angus, "Saving in Developing Countries: Theory and Review", *World Bank Economic Review*, Special Issue, Proceedings of the Annual World Bank Conference on Development Economics.

______, "Saving and Liquidity Constraints", *Econometrica*, 59(5)

Deaton, Angus and Margaret Grosh, "The Consumption Model in the Living Standards Measurement Study", manuscript, May 1997

Deolalikar, Anil.B. "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India", *Review of Economics and Statistics*, 70(3)

Dor,-Avi, Paul Gertler and Jacques van der Gaag, "Non-price Rationing and the Choice of Medical Care Providers in Rural Cote d'Ivoire", *Journal of Health Economics*, 6(4), December 1987, pages 291-304.

Eswaran, Mukesh and Ashok Kotwal, "Credit as Insurance in Agrarian Economies", *Journal of Development Economics*, 31(1), pp. 37-53

Gersovitz, Mark, "Savings and Development", Chapter 10, *Handbook of Development Economics*, vol. I.

Greenwood, Jeremy, and Boyan Javanovic, "Financial Development, Growth and the Distribution of Income," *Journal of Political Economy*, October 1990, 98, 93-118.

Greene, William H. "Econometric Analysis", 2nd ed., Prentice Hall, 1993.

Grossman, Michael, "On the Concept of Health Capital and the Demand for Health" *Journal of Political Economy*, Vol. 80, No. 2 (March-April 1972), 223-55.

Hsiao, Cheng, "Analysis of Panel Data", Cambridge University Press, 1986.

Kochar, Anjini "Explaining Household Vulnerability to Idiosyncratic Shocks", *American Economic Review*, Vol. 85, No. 2 (May 1995), 159-64.

______, "Smoothing Consumption by Smoothing Income: Hours-of-Word Responses to Idiosyncratic Agricultural Shocks in Rural India", *The Review of Economics and Statistics*, 1999, Vol. 81, No. 1 (February 1999), 50-61.

Lim, Youngjae, and Robert Townsend, "Currency Transaction Patterns, and Consumption Smoothing Theory and Measurement in ICRISAT Villages," working paper, University of Chicago, 1994.

Lipton, Michael and Martin Ravallion, "Poverty and Policy", Chapter 41, *Handbook of Development Economics*, Vol. II.

Lucas, Robert and Oded Stark, "Motivations to Remit: Evidence from Botswana", *Journal of Political Economy*, 93(5)

Morduch, Jonathan, "Income Smoothing and Consumption Smoothing", Journal of Economic Perspective, Summer 1995

Paulson, Anna, "Insurance Motives for Migration: Evidence from Thailand", dissertation, University of Chicago, 1994.

Paxson, Christina, and C. Chaudhuri, "Consumption Smoothing and Income Seasonality in Rural India", manuscript, 1994.

Paxon, Christina, "Consumption and Income Seasonality in Thailand", *Journal of Political Economy*, Vol.101, No. 11 ()

Picone, Gabriel, Matine Uribe, and R. Mark Wilson, "The Effect of Uncertainty on the Demand for Medical Care, Health Capital and Wealth", *Journal of Health Economics*, 17

Ray, Debraj, "Development Economics", Princeton University Press, 1998.

Rose, Elaina, "Consumption Smoothing and Excess Female Mortality in Rural India," manuscript, University of Washington, 1994.

Rosenzweig, Mark and Oded Stark, "Consumption Smoothing, Migration and Marriage: Evidence From Rural India", *Journal of Political Economy*, 97(4)

Rosenzweig, Mark, "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries", *The Economic Journal*, December 1988, 98(393)

Rosenzweig, Mark and Kenneth Wolpin, "Credit Market Constraints, Consumption Smoothing and the Accumulation of Durable Assets in Low-Income Countries: Investments in Bullocks in India", *Journal of Political Economy*, 101(1)

Strauss, John and Duncan Thomas, "Prices, infrastructure, household characteristics and child height", *Journal of Development Economics*, 39(2)

______, "Human Resources: Empirical Modeling of Household and Family Decisions", Chapter 34, *Handbook of Development Economics*, vol. II

Townsend, Robert "Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies", *Journal of Economic Perspectives*, Summer 1995, pp 83-102

Udry, Christopher, "Credit Markets in Northern Nigeria: Credit as Insurance in a Rural Economy", *World Bank Economic Review*, 1991, 4:3, 251-71.

Viravaidya, and Mechai Stasia Obremskey and Charles Myers, "The Economic Impact of AIDS on Thailand", manuscript, 1991

CURRICULUM VITAE

RONG YI

1156 Commonwealth Ave.
Apt. 51,
Allston, MA 02134

Tel: (617) 264-9808
Fax: (617) 353-4449
Email: rongyi@bu.edu

EDUCATION

***Boston University, Department of Economics, Boston, MA**, 1995-present*
- Ph.D. in Economics (expected in Jan. 2001)
- Dissertation Title: "Risks, Insurance, and the Loss of Work due to Illness among Low-Income Households."

***Boston University, Department of Economics, Boston, MA**, September 2000*
- Master of Arts in Political Economy

***Peking University, Beijing, China**, July 1995*
- Bachelor of Science in Regional Economics

***Beijing Film Academy, Beijing, China**, July 1995*
- Minor in Screenwriting

SPECIALIZATION

- Applied Econometrics
- Applied Industrial Organization / Insurance
- Economic Development

RESEARCH PAPERS & PUBLICATIONS

- "Risks, Insurance, and the Decision to Take Sickness Leave: Townsend Revisit with Sample Selection in General Equilibrium", working paper, Department of Economics, Boston University, 2000.
- "Sample Selection in the ICRISAT Labor Data", working paper, Department of Economics, Boston University, 2000.
- "Full Insurance under Nonseparable Preferences", working paper, Department of Economics, Boston University, 1999.

- "On the Tolerance of City—A Structuralist Approach to Urban Space Organization", mimeo, Peking University, Beijing, China, 1995.
- "Real Estate Market in an Economy in Transition", *Real Estate Development and Management*, Peking University Press, Beijing, China, 1995.
- "Real Estate Financing", *Real Estate Development and Management*, Peking University Press, Beijing, China, 1995.
- "Introduction to Foreign Real Estate Market Operations", *Real Estate Development and Management*, Peking University Press, Beijing, China, 1995.
- "Land Use Analysis and Future Development of Peking University Campus Estate", working paper, Campus Design Project, Peking University, Beijing, China, summer 1994.
- "The Impact of Northeastern Railway Construction on Regional Development", working paper, Changchun Group Incorporated, Beijing, China, spring 1994.
- "Taiwan Investment in Mainland China", *Journal of Information and Economics*, Guangxi, China, November 1993.

RESEARCH EXPERIENCE

Research Assistant
Department of Economics, Boston University, Spring 2000

- Cleaned and restructured cross-sectional time-series data from Italy with SAS and STATA for Prof. Maristella Botticini and Prof. Aldo Rustichini; drafted progress reports.

Research Fellow
Department of Economics, Boston University, Fall 1997 to Spring 1998

- Reviewed leading economic and financial journals and conducted literature research on health care, telecommunications, information technology and financial industries;
- Designed and maintained the homepage for the Industry Study Program.

Research Assistant
School of Law, Boston University, Summer 1996

- Reviewed leading economic journals and prepared article summaries on predatory pricing and other antitrust issues for Prof. Joseph Brodley.

TEACHING EXPERIENCE

Part-time Faculty
School of Management, College of Arts and Sciences & Metropolitan College

Boston University, Summer 1998 to present.

- Taught full courses independently to undergraduate students in regular semesters and summer terms, prepared lectures, auxiliary readings and examinations, and invited guest speakers on various occasions;
- Received high student evaluations on teaching skills;
- Courses taught: Money & Capital Markets, Money and Banking, Intermediate Macroeconomics, Development Economics, Introduction to Macroeconomics, Introduction to Microeconomics, and General Economics.

Teaching Assistant
College of Arts and Sciences, Boston University, Fall 1995 to Spring 1998.

- Led discussion and review sessions for Introductory Microeconomics and Introductory Macroeconomics, graded homework and examinations.

SKILLS

- Strong writing, communication and presentation skills;
- Strong economic modeling and econometric application skills;
- SAS, STATA and GAUSS programming;
- Experienced with handling large cross-sectional and time-series data sets;
- Proficient with Scientific Workplace and Microsoft Word, Excel and PowerPoint.

AWARDS

- Teaching Fellowship, Boston University, 1995-1998.
- Industry Study Program Research Fellowship, Boston University, 1997-1998.