THREE ESSAYS ON THE ECONOMICS OF HEALTH IN DEVELOPING COUNTRIES

by

Eiji Mangyo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2005

Doctoral Committee:

Associate Professor Albert Francis Park, Chair
Professor Charles C. Brown
Associate Professor Michael E. Chernew
Assistant Professor Andrew M. Coleman

UMI Number: 3186700

To our parents

ACKNOWLEDGEMENTS

There are several people that I am indebted to not only for their highly professional expertise that was indispensable for the completion of this dissertation, but also for their personalities that helped me through the Ph.D. program. First and foremost, Albert Park my supervisor and mentor. Through his fine example, I learnt what it was to be a university professor. Throughout my academic life in Michigan, he continued to advise me constructively and patiently. Before entering the Ph.D. program, I had no idea as to how to conduct academic research. I can now address a research question with significant importance using reasonably scientific methods, largely due to Albert Park's patient instruction. This dissertation benefits a lot from his insight and communication skills. He sees the essence of the matter and explains it in clearly understandable terms. His sincerity toward both economic research and toward nurturing students has set an inspirational model for my career.

Charlie Brown, Mike Chernew, and Andrew Coleman repeatedly read my drafts and provided useful suggestions. They were always so supportive that I was able to continue research in a very comfortable setting. Charlie Brown also emotionally helped me when my research did not make progress. I also thank Bob Stern for his extremely generous support when I was a master student in public policy. He encouraged me to study at the Ph.D. level and inspired my interest in academic research. Stan Panis gave me clear instructions about aML programming that are based on his deep economic insight. Hector Chade helped me to pursue a theoretical model that stimulated subsequent empirical research. Catherine Cross provided helpful and sincere responses to my data inquires.

Finally but not the least, I am grateful to my wife, Kyoko, and our parents for continuous support during my academic endeavor. Our parents often shipped us a box of Japanese food as well as caring mind. Because I have not been able to express my sincere

gratitude to our parents, I would like to dedicate this dissertation to our parents as an indication of my heartfelt acknowledgment.

TABLE OF CONTENTS

DEDICATION..........ii

ACKNOWLEDGEMENTS..........iii

LIST OF TABLES..........viii

LIST OF FIGURES..........x

LIST OF APPENDICES..........xi

Chapter 1 Who Benefits More from Higher Household Consumption? The Intra-household Allocation of Nutrients in China..........1

- **1. Introduction**..........1
- **2. Previous Research on Vulnerable Demographic Groups within Households**..........4
- **3. Theory**..........5
- **4. Data**..........11
- **5. Econometric Analyses**..........12
  - 5.1 Econometric Model..........12
  - 5.2 Instruments..........16
  - 5.3 Estimation Results..........17
  - 5.4 Controlling for Work Hours..........22
  - 5.5 Do Results Differ for the Poor?..........23
- **6. Demographic Differences in Nutrient Sufficiency**..........24
- **7. Conclusions**..........26
- **Appendix 1.1 Theory Appendix**..........39
- **Appendix 1.2 Full Results of 2SLS Regressions**..........47
- **References**..........69

Chapter 2 Measuring the Impact of Health Insurance on Physician Visits by the Elderly: A Natural Experiment in Taiwan......73

1. **Introduction**......73
2. **Previous Research**......77
3. **Data**......78
4. **Descriptive Evidence**......79
5. **Estimation Strategies and Results**......82
   - 5.1 Two-Part Model......82
   - 5.2 Fixed Effects Models......83
     - 5.2.1 Physician visits without distinction between any visit and conditional visits......83
     - 5.2.2 Any visit to a physician......86
     - 5.2.3 Physician visits conditional on any visit......87
     - 5.2.4 Discussion......89
   - 5.3 Random Effects Model with Endogenous Health Insurance......91
     - 5.3.1 Econometric model......91
     - 5.3.2 Identification issues......93
     - 5.3.3 Estimation results......94
6. **Conclusions**......97

Appendix 2.1 Explanation of the Negative Binomial Model (Conditional-Visit Equation) in the RE Simultaneous-Equation System......118

References......120

Chapter 3 The Effect of Water Accessibility on Child Health in China......122

1. **Introduction**......122
2. **Data**......126
3. **Econometric Analyses**......127
   - 3.1 Base-line Child-Specific Fixed-Effects Model......127
   - 3.2 Addressing Potential Endogeneity of Household Income......129
   - 3.3 Addressing Potential Endogeneity of Access to Clean Water......131
   - 3.4 Parental Education and Water Access......132

3.5 Non-Random Project Placement....................................133
3.6 Our Identification Strategies Effective?.............................136
4. Conclusions..138
Appendix 3.1 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects...........................155
References..156

LIST OF TABLES

1.1 Whose Nutrient Intake Is More Income Elastic? (The Case with Non-Linear Production Functions)....28

1.2 Summary Statistics of Multiple-Person Households by Income Group....29

1.3 Elasticity Estimates of Nutritional Intakes by Demographic Group....30

1.4 Elasticity Estimates of Nutritional Intakes by Demographic Group (Farm Households Only)....31

1.5 Pair-wise 3SLS Results on Elasticities of Protein Intakes by Demographic Group....32

1.6 Pair-wise 3SLS Results on Elasticities of Caloric Intakes by Demographic Group....32

1.7 Elasticity Estimates of Nutritional Intakes by Demographic Group with Endogenous Labor Supply....33

1.8 Elasticities of Protein Intake by Demographic Group with Endogenous Labor Supply (Poor and Richer Households Separately)....34

1.9 Elasticities of Caloric Intake by Demographic Group with Endogenous Labor Supply (Poor and Richer Households Separately)....35

1.10 Chinese Dietary Reference Intakes (DRIs) for Calories and Proteins....36

1.11 2SLS Regressions of Individual Protein Intakes (Full Results) (A) Main-Equation Results....47

1.12 2SLS Regressions of Individual Protein Intakes (Full Results) (B) First-Stage Results....52

1.13 2SLS Regressions of Individual Caloric Intakes (Full Results) (A) Main-Equation Results....58

1.14 2SLS Regressions of Individual Caloric Intakes (Full Results) (B) First-Stage Results....63

2.1 Trend in Medical Utilization in Taiwan....98

2.2 Health Related Indicators for Selected Countries....99

2.3 Comparison of Age and Gender Distribution between Completed Sample and Population....101

2.4 Mean Comparison of Demographic and Financial Characteristics between the Previously Insured and the Previously Uninsured....102

2.5 Changes in Physician Visits by Health Insurance Status prior to NHI....103

2.6 Mean Changes in Physician Visits between 1989 and 1993 by Health Insurance Status....104

2.7 Changes in Health Insurance Status from 1989 to 1993 104

2.8 Fixed-Effects Linear Regression Results and Fixed-Effects Poisson Regression Results for Physician Visits 105

2.9 Fixed-Effects Logit Regression Results for Any Visit 107

2.10 Fixed-Effects Poisson Regression Results for Conditional Visits 108

2.11 Supply-Side Indicators of Medical Services 109

2.12 Occupations with Access to Health Insurance to Spouses before 1995 110

2.13 Regression Results of Simultaneous-Equation System 111

3.1 Description of Variables 141

3.2 Summary Statistics of Child/Household/Community Characteristics 142

3.3 Sample Children by "Near Water" Status in 1991 and 1993 143

3.4 Cross-Tabulation of Changes in Water Access and Changes in BMI z Score 143

3.5 Econometric Results 144

3.6 Number of Sample Households within Communities 148

3.7 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria 148

3.8 Econometric Results with Community Fixed Effects 149

3.9 Community-level Dynamic Correlation between Changes in % Households with Access to "Near Water" and Changes in Various Community Characteristics 152

3.10 Probit Regressions of Treatment (Changes in Water Access) 153

3.11 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects (Children Age 15 or Younger) 155

3.12 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects (Children Age 10 or Younger) 155

LIST OF FIGURES

1.1 Demographic Difference in Protein SI..37

1.2 Demographic Difference in Caloric SI..38

2.1 National Trend of Outpatient Visits 115

2.2 National Trend of Inpatient Visits..115

2.3 National Trend of Emergency Visits...115

2.4 Sample Trend of Physician Visits...116

2.5 Sample Difference in Physician Visits...116

2.6 Sample Trend of Any Visit...116

2.7 Sample Difference in Any Visit..116

2.8 Sample Trend of Conditional Visits...116

2.9 Sample Difference in Conditional Visits...116

2.10 National Trend of Outpatient Visits per Physician....................................117

LIST OF APPENDICES

1.1 Theory Appendix..39

1.2 Full Results of 2SLS Regressions..47

2.1 Explanation of the Negative Binomial Model (Conditional-Visit Equation) in the RE Simultaneous-Equation System..118

3.1 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects..155

Chapter 1

Who Benefits More from Higher Household Consumption? The Intra-household Allocation of Nutrients in China

Abstract

Previous studies find that human capital investments in boys are less income elastic than investments in girls, attributing this result to favoritism toward boys. I show theoretically that it is plausible for more productive or favored household members to have higher income elasticities. I then investigate this question empirically, utilizing panel data on individual nutrient intake from the China Health and Nutrition Survey (CHNS) to analyze how changes in household per-capita nutrient intake affect the intra-household allocation of nutrients. To deal with potential biases due to omitted variables and simultaneity, I use measures of rainfall variation as instruments. I find that nutritional intakes are more elastic for males (especially prime-age men) than for females, and significantly less elastic for the elderly.

1. Introduction

The intra-household allocation of resources is an important economic issue particularly from the development standpoint as the welfare of individuals can vary drastically between seemingly similar households due to how resources are allocated within the household unit. Demographic differences in the income and price elasticities of consumption and human capital investments provide policy makers with essential information to evaluate the welfare impacts of income generation and pricing policies. Previous studies have found that investments in favored demographic groups are less price and income elastic than investments in less favored demographic groups.[1] For

[1] The literature is most abundant in education with particular interest in gender differences in schooling. For price elasticities, see Schultz 1987, King and Lillard 1987, de Tray 1988, Gertler and Glewwe 1992,

example, Alderman and Gertler (1997) propose a theoretical model in which human capital investments are less income and price elastic for favored children than for less favored children, and empirically show that the demand for medical care is more price and income elastic for girls than for boys in Pakistan. Behrman and Deolalikar (1990) find that nutrient intakes for females systematically have more negative price elasticities than those for males, which they conclude may leave females particularly vulnerable during times of food shortages. Behrman (1988) also finds that girls' nutrition suffers more than boys' nutrition in the lean agricultural season when household resources are often depleted. Rose (1999) finds that favorable rainfall shocks increase the probability that girls will survive more than the probability boys will survive in rural India. Dercon and Krishnan (2000) find that BMI decreases in response to unpredicted illness shocks are larger for women than for men in poor Ethiopian households.

Unfortunately, the previous theoretical model is highly stylized and is largely influenced by the idea that human capital investments in boys should be more of a necessity than those in girls, thus the income elasticity must be lower for boys than for girls.[2] Empirical studies (both cross-sectional and panel studies) on the intra-household allocation of resources, including those mentioned above, do not adequately control for potential confounding factors correlated with human capital investments.[3] This chapter examines how the intra-household allocation of nutrition responds to changes in household consumption levels, and challenges the theoretical and empirical findings of the existing literature. First, I present a theoretical model that demonstrates that it is inconclusive whether a more productive member (or more favored member) has a greater

---

Levy 1996, Sipahimalani 1999, and World Bank 2001. For income elasticities, see Schultz 1987, and de Tray 1988.

[2] The theoretical model Alderman et al (1997) propose is a two-period model where the parents maximize their utility of the following form: $U = F(C_1) + G(C_2, W_b, W_g)$. C_1 and C_2 are the parents' consumption in the first and second periods, and W_b and W_g are the boy's and girl's wealth in the second period. W_b and W_g are assumed the linear functions of human capital investments made in the first period: $W_b = bH_b$ and $W_g = gH_g$. C_2 is assumed the linear function of the boy's and girl's wealth: $C_2 = \beta W_b + \tau W_g$, where b, g, β, and τ are all constants. Alderman et al (1997) assume away the case where $|\partial^2 G / \partial H_b \partial H_b| < |\partial G / \partial H_g \partial H_g|$, which leads to unambiguously more elastic investments in the girl than in the boy with respect to household income and the price of human capital investments.

[3] Dercon and Krishnan (2000) seriously consider the problem of potential confounding factors correlated with unpredicted illness shocks (such as reverse causality from BMI changes to illness shocks), but they do not address the problem econometrically.

or smaller elasticity of nutritional intake with respect to income. Households are concerned about both equity and efficiency when allocating resources among household members. As household food consumption increases, increases in the allocations to individuals will depend upon how fast marginal utilities and productivities of individual members fall relative to other household members.

Despite the difficulty of deriving strong theoretical predictions, measuring differences in income elasticities among demographic groups is still an important empirical question, because it sheds light on the welfare consequences of policies that affect income and consumption levels. In the empirical section of the chapter, using data from China I estimate how the nutritional intakes of individuals from different demographic groups respond to changes in total household food consumption. We are particularly interested in the relative magnitudes of the nutrient-intake elasticity among six demographic groups: prime-age men, prime-age women, elderly men, elderly women, boys, and girls. Previous empirical evidence and numerous anecdotes suggest that males are more favored than females in China. Especially, prime-age men are considered the most favored and productive demographic group in the society. This chapter examines whether the nutrient-intake elasticity with respect to total household food consumption is lower for males than for females, and for prime-age men than for other demographic groups in accordance with the findings of previous research. To deal with potential biases due to omitted variables and simultaneity, I use measures of rainfall variation as instrumental variables. As far as I know, this is the first panel study that controls for inter-temporal confounding factors in examining the effect of household wealth on intra-household allocation issues.

There are several advantages of focusing on nutrient allocation in examining intra-household decision-making. First, food is the major consumption expenditure of households in most developing countries. In 1993, the last year of survey data used in this study, expenditures on food accounted for about 50% of total expenditures in urban China and about 60% in rural China. Moreover, nutrients are not easily substitutable by other goods, making it unlikely for allocations of other goods to compensate for inequities in food resource allocation. Finally, as the consumption item most essential for survival, a focus on food allocation may highlight the tradeoffs between equity and efficiency concerns (Pitt et al 1990).

The rest of the chapter is organized as follows. Section 2 briefly reviews the relevant literature on the intra-household allocation of resources. Section 3 presents a theoretical model of the intra-household allocation of nutrients. Section 4 discusses the data used in the empirical analyses. It also provides descriptive statistics of the sample households. Section 5 presents the results of econometric analyses, and Section 6 concludes.

2. Previous Research on Vulnerable Demographic Groups within Households

Later, this chapter theoretically shows that elasticities of human capital investments tell us nothing about who is favored or more productive within households, and empirically calculates that the nutrient-intake elasticity with respect to total household consumption of food is highest for prime-age men. To interpret our empirical results using our theoretical results, we need some evidence about who in households are more favored or productive. One objective of this section is to look at some evidence about which demographic groups are less favored or productive within households. Another purpose is to see why the intra-household allocation of resources is important in China.

Using household data from rural Pakistan, Kochar (1999a) shows that an individual's predicted wage positively affects the amount of medical expenditures on that individual. This implies that the elderly are at a disadvantage in terms of medical attention when they are ill, because their earning ability attenuates as they age. Miguel (2003) shows that extreme rainfall (drought or flood) in rural Tanzania leads to a large increase in murders of "witches"- typically elderly women killed by relatives - but not other murders. These results are consistent with a household allocation model that puts emphasis on productivity as the main determinant in the intra-household allocation of resources.

China is a country for which there have been few previous studies on the intra-household allocation of resources. Traditionally, China is considered to be a pro-male society. Chinese households have been characterized as paternalistic, and descendants in the male line have carried family names (Lee and Wang 1999). The reported sex ratio (of boys to girls) at birth reached 116.9 in 2000 (Wiseman 2002), and this has been widely interpreted as reflecting favoritism toward boys ("missing girls" in China; See Coale 1991, Johansson and Nygren 1991, Zeng et al 1993, and Junhong 2001). Park and Rukumnuaykit (2004) find that rural fathers reduce nutrient intakes more if they reside with boys than if they reside with girls after controlling for other relevant factors influencing fathers' nutrient intakes, which is consistent with favoritism toward boys. Yu and Sarri (1997) show that disparity in health between men and women in China narrowed but still existed in 1990 (in terms of the Physical Quality of Life Index: PQLI) and that China and South Asian countries were behind other Asian countries in gender equality in the early 1990s (in terms of the Gender-Related Development Index: GDI).

Understanding intra-household allocations of resources is important for evaluating the welfare of the elderly in China. The dominant form of living arrangement for the elderly is living with adult children (66% in urban areas and 73% in rural areas, Lee et al 1998), and the vast majority of the Chinese elderly rely on financial support from adult children. In 1992, only 5.7% of the elderly had pension income in rural China where more than three-quarters of the elderly reside (Pei et al 1999). Moreover, China's old age population is predicted to grow rapidly both in absolute number and as the ratio of total population.[4]

Understanding how resources are allocated to children also has significant policy relevance in China. Child malnutrition among both pre-schoolers and school children was substantial in the early 1990s.[5] Even mild to moderate malnutrition is associated with greater mortality (Pelletier et al 2002; Schroeder et al 1997) and poor school performance

---

[4] The Chinese census of 1990 indicates that those aged 60 and older were about 97 million in size and comprised 8.6% of the total population. By 2025, they will increase to about 264-298 million (which is larger than the total population of the United States) and constitute 17-19% of the population (Kwong et al 1992).

[5] According to the WHO Global Database on Child Growth and Malnutrition (www.who.int/nutgrowthdb/, accessed on July 31, 2004), substantial proportions of Chinese children under age five suffered from stunting (rural 34.5%, urban 18.2%), underweight (rural 19.2%, urban 9.4%), and wasting (rural 3.6%, urban 2.6%) in 1992. Morgan (2000) provides useful statistics regarding malnutrition among school children in China. The mean height-for-age of eleven-year-old children in rural and urban China in 1995 approximately stood at the 20^{th} and 40^{th} percentiles, using the height for age of US children as the reference population. The mean height-for-age of 7-year-old children in rural and urban China in 1995 fell at the 22^{nd}-23^{rd} and the 40^{th}-42^{nd} percentiles, respectively. The mean height-for-age of 17-year-old children in rural and urban China in the same year fell at the 11^{th}-16^{th} and the 21^{st}-26^{th} percentiles, respectively.

(Jamison 1986). Effective policies to improve child nutrition need to take into account the household dynamics of food allocation.

3. Theory

In this section, we develop a simple model to show how exogenous increases in total household resources affect intra-household allocations among individual members with different productivities and/or different weights in the household's utility function. Consider a household that contains one member with higher earnings potential (e.g. a prime-age adult) and one member with lower earnings potential (e.g. a dependent such as an elderly member or child). The household is concerned about both efficiency and equity (Pitt et al 1990), and solves the following utility maximization problem:

$$\begin{gathered} \underset{X_p, X_d}{Max}\ U(X_p, X_d) \\ \text{s.t. } X_p + X_d = Y(X_p, X_d, Y_0) \end{gathered} \tag{1}$$

Household utility is a function of the nutritional intakes of the prime-age member X_p and the dependent member X_d, respectively.[6] Since nutrition also affects productivity, total income Y is a function of X_p and X_d, as well as exogenous income Y_0. Natural assumptions are that the utility and production functions are increasing and concave in each member's nutritional intake ($\frac{\partial U}{\partial X_p}, \frac{\partial U}{\partial X_d} > 0$, $\frac{\partial^2 U}{\partial X_p \partial X_p}, \frac{\partial^2 U}{\partial X_d \partial X_d} < 0$ for the utility function and $\frac{\partial Y}{\partial X_p}, \frac{\partial Y}{\partial X_d} > 0$, $\frac{\partial^2 Y}{\partial X_p \partial X_p}, \frac{\partial^2 Y}{\partial X_d \partial X_d} \leq 0$ for the production function). Further, there is a difference in productivity between the prime-age and dependent members ($\frac{\partial Y}{\partial X_d} \leq \frac{\partial Y}{\partial X_p}$ when $X_p = X_d$). To ensure that a unique solution exists, we also assume $0 \leq \frac{\partial Y}{\partial X_d}, \frac{\partial Y}{\partial X_p} < 1$ at the optimum.

[6] Adding another purchased good to the model does not alter the main qualitative results.

Of interest here are the nutrient-intake elasticities of the prime-age and dependent members with respect to exogenous household income Y_0. Under fairly general assumptions, it is easy to show that both members' nutritional intakes are increasing in exogenous income Y_0 ($\frac{dX_p^*}{dY_0} > 0$ and $\frac{dX_d^*}{dY_0} > 0$), thus, the nutrient-intake elasticities of both members with respect to Y_0 are positive ($\eta_m = \frac{dX_m^*}{dY_0}\frac{Y_0}{X_m^*} > 0$ for $m = p, d$).[7] However, deriving the relative magnitude of the two elasticities is more complicated.

To make the analysis tractable, we make several simplifying assumptions.

Assumption 1: The household utility function is separable in X_p and X_d (i.e. $U_{X_p X_d} = 0$), and individuals share a common utility function $u(X)$.

Under Assumption 1, the household utility function simplifies to $U(X_p, X_d) = u(X_p) + \beta u(X_d)$ where $u' > 0$ and $u'' < 0$. β captures favoritism, where equal treatment implies $\beta = 1$.

Assumption 2: The household production function is separable in X_p and X_d (i.e. $Y_{X_p X_d} = 0$), and individual production functions differ only by a multiplicative constant.

Under Assumption 2, the household production function simplifies to $Y(X_p, X_d, Y_0) = \omega_p y(X_p) + \omega_d y(X_d) + Y_0$ where $y' > 0$ and $y'' \le 0$. ω_p and ω_d capture productivities of the members, where a greater productivity of the prime-age member implies $\omega_p > \omega_d$. For the moment, we assume that the individual production functions are linear for both members.

Assumption 3: The individual production functions are linear (i.e. $y'' = 0$).

[7] See Appendix 1.1 for a proof of this claim.

Under Assumptions 1, 2, and 3, the maximization problem the household solves is analogous to a typical consumer demand problem where there are two goods X_p and X_d whose prices are $1-\omega_p$ and $1-\omega_d$, respectively.

$$\begin{aligned} &\underset{X_p, X_d}{Max}\ U(X_p, X_d) = u(X_p) + \beta u(X_d) \\ \text{s.t.}\quad &(1-\omega_p)X_p + (1-\omega_d)X_d = Y_0 \end{aligned} \tag{2}$$

Here, it is convenient to introduce the concepts of necessities and luxuries from consumer demand theory. According to Tolley and Gieseman (1963),

> "ρ_i is the percentage change in marginal utility of the ith good associated with a one per cent increase in that good. A larger negative value of ρ_i indicates that the good is a necessity, and a small negative value indicates that it is a luxury" (p. 500).

For our problem, we can define ρ as a function of nutrient consumption X:

$$\rho(X) = \frac{du'(X)/u'(X)}{dX/X} = \frac{u''(X)X}{u'(X)} \tag{3}$$

Thus, $\rho(X)$ evaluated at a particular equilibrium (X_p^*, X_d^*) may differ for prime-age and dependent members, depending on their levels of consumption. From this definition of $\rho(X)$, we see that $R(X) = -\rho(X) > 0$ where $R(X)$ measures relative risk aversion.[8] If $R(X)|_{X=X_p^*} > R(X)|_{X=X_d^*}$ for any equilibrium (X_p^*, X_d^*), then X_p^* is more of a necessity than X_d^* and has a lower income elasticity. If $R(X)|_{X=X_p^*} < R(X)|_{X=X_d^*}$ for any equilibrium (X_p^*, X_d^*), then X_p^* is more of a luxury than X_d^* and has a higher income elasticity.

[8] Essentially, $R(X)$ is a measure of the curvature of the utility function although it is commonly called relative risk aversion.

It is easy to show that if the prime-age member is more productive ($\omega_p > \omega_d$) and/or more favored ($\beta < 1$), then $X_p^* > X_d^*$. Then, it follows that the relative elasticities of nutrient intakes by the prime-age and dependent members will depend on whether $R(X)$ increases or decreases in X.

Proposition 1: Under Assumptions 1, 2, 3, and $X_p^* > X_d^*$ for any equilibrium, $\frac{dR(X)}{dX} > 0$ for any X implies $\eta_d > \eta_p$, $\frac{dR(X)}{dX} = 0$ for any X implies $\eta_d = \eta_p$, and $\frac{dR(X)}{dX} < 0$ for any X implies $\eta_d < \eta_p$.

(Proof) See Appendix 1.1.

Thus, whether prime-age or dependent members have higher elasticities depends on the concavity of the individual utility function. If relative risk aversion decreases in X, then it is possible for the income elasticity of nutrient intake to be higher for the prime-age member than for the dependent member.

Next, we consider a case where there is favoritism toward the prime-age member ($\beta < 1$). We present the following optimality condition derived from the first order conditions:

$$\frac{1-\omega_d}{1-\omega_p} = \frac{\beta u'(X_d)}{u'(X_p)} \tag{4}$$

This equality simply equates the marginal rate of transformation and the marginal rate of substitution of nutrient intakes by the members. The numerator (denominator) on the left-hand side of (4) can be thought of as the shadow price of nutrient intake by the dependent (prime-age) member. The numerator (denominator) on the right-hand side is the marginal utility of the nutrient to the dependent (prime-age) member. In equilibrium, the ratio of the shadow prices must be equal to the ratio of the marginal utilities. If we divide both sides of the equality (4) by $\beta < 1$, we can immediately see that favoritism toward the prime-age member is equivalent to lowering the relative price of consumption by the prime-age member (or increasing the prime-age member's productivity). Thus, favoritism increases the level of consumption of the favored member but does not affect our theoretical results on differences in the elasticity of consumption, which depend on the curvature of the utility functions.

Next, I relax the linearity assumption for the production function (Assumption 3). It is more plausible that the production function linking nutrient intake and output is concave rather than linear. Under Assumptions 1 and 2, the household maximization problem simplifies to the following problem:

$$\begin{aligned} &\underset{X_p, X_d}{Max}\ U(X_p, X_d) = u(X_p) + \beta u(X_d) \\ \text{s.t.} \quad & X_p + X_d = \omega_p y(X_p) + \omega_d y(X_d) + Y_0 \end{aligned} \tag{5}$$

The optimality condition under Assumptions 1 and 2 is as follows:

$$\frac{1 - \omega_d y'(X_d)}{1 - \omega_p y'(X_p)} = \frac{\beta u'(X_d)}{u'(X_p)} \tag{6}$$

In contrast to the linear production case, the ratio of the shadow prices (left-hand side) varies as X_p and X_d change. To deal with this case, it is helpful to introduce the analogous concepts to necessities and luxuries for the production function.

$$Q(X) = -\frac{dy'(X)/y'(X)}{dX/X} = -\frac{y''(X)X}{y'(X)} \tag{7}$$

One can show that under Assumptions 1, 2, and $X_p^* > X_d^*$ for any equilibrium, the sign of $\eta_d - \eta_p$ is determined by the signs of both $\frac{dR(X)}{dX}$ and $\frac{dQ(X)}{dX}$ as well as other conditions, as depicted in Table 1.1 (See Appendix 1.1). We consider three cases. First, productivity in equilibrium is always equal for the prime-age and dependent members ($Y_{X_p} = Y_{X_d}$). Second, productivity in equilibrium is always greater for the prime-age member than for the dependent ($Y_{X_p} > Y_{X_d}$). Finally, productivity in equilibrium is always greater for the dependent member ($Y_{X_p} < Y_{X_d}$).

In all three sub-tables in Table 1.1, moving toward the northwest direction motivates the household to deliver nutrients proportionally more toward the dependent than toward the prime-age member when household exogenous income increases, resulting in a higher income elasticity of nutrient intake for the dependent. Moving toward the southeast direction in the tables gives the opposite motivation to the household, resulting in a higher income elasticity of nutrient intake for the prime-age member. The theory shows that the shapes of the utility and production functions determine whose nutrient intake is more income elastic, and it is equally plausible for the more productive or favored member to have a higher income elasticity of nutrient intake in comparison with the dependent member. This is rather an empirical question.

4. Data

Data from the second (1991) and third (1993) waves of the China Health and Nutrition Survey (CHNS) are used for the analyses.[9] The CHNS is one of the few datasets from developing countries that has information on individual nutrient intake for all household members over time, making it particularly well-suited for examining intra-household resources allocation decisions.

Each wave of the CHNS consists of a household survey, individual surveys of health and nutrition, an elderly survey, an ever-married women survey, a community survey, and a health and family planning facility survey. The survey population is drawn from eight of China's thirty-one provinces, located throughout the country: Guangxi, Guizhou, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. A multistage, random cluster approach was used to construct the sample in each of the eight provinces. The 190 primary sampling units consisted of 32 urban neighborhoods, 30 suburban neighborhoods, 32 towns, and 96 villages. The household survey includes information on household income and assets, as well as time allocation by household members.

[9] Complete data on individual nutrient intake are not available in the first (1989) and fourth (1997) waves.

The CHNS is notable for the high quality of its health and nutrition data. In 1991 and 1993, individual dietary intake for three consecutive days was enumerated for all individuals in each surveyed household. Individuals were asked each day to report all food consumed away from home and at home on a 24-hour recall basis. Household food consumption was verified by measuring changes in food inventories from the beginning to the end of each day. All processed foods (including edible oils and salt) were measured at the beginning of the survey period. All purchases, home production, and processing foods were recorded. Whenever foods were brought into the household unit, they were weighed. Preparation waste (e.g., spoiled rice, discarded cooked meals fed to pets or animals) was estimated when weighing was not possible. At the end of the survey, all remaining foods were again weighed and recorded. The number of household members and visitors present at each meal was recorded.[10]

In Table 1.2, I report summary statistics of multiple-person households by income group, using the pooled sample households from 1991 and 1993. In the lowest income group, the vast majority of households are farmers, while in the top income group, a little less than half households are farmers. Richer households contain fewer children on average in comparison with poorer households. Poorer households tend to live in village areas and richer households tend to live in non-village areas. Provinces of residence also differ significantly for different income groups. For example, 19% of poor sample households come from Henan while less than six percent of (relatively) rich households come from the same province. In a similar vein, Jiangsu is the province of residence for 18% of high-income households, while only nine percent of non-rich households come from Jiangsu. Household size is slightly smaller for richer households than for poorer households.

5. Econometric Analyses

5.1 Econometric Model

To estimate the response of intra-household nutrient allocation to changes in total household food consumption, I first estimate the average response separately for each of

---

[10] Further information of the CHNS is available at http://www.cpc.unc.edu/projects/china (accessed on December 11, 2004).

six demographic groups based on gender and age in 1991: prime-age men (between 18 and 59 years old), prime-age women, elderly men (age 60 or older), elderly women, male children (age 17 or younger), and female children. The main estimating equation is the following:

$$\log(N_{ijkt}) = \alpha_{im}W_{jkt} + \beta_{im}\log(y_{ikt}) + \gamma_{im}X_{kt} + \varepsilon_{ijk} + \mu_{iat} + \mu_{ist} + \mu_{irt} + \tau_{ijkt} \qquad (8)$$

where i, j, k, t, and m index nutrient, household member, household, time, and demographic group, respectively. N_{ijkt} is daily intake of nutrient i consumed by household member j in household k at time t; y_{ikt} is per-capita nutrient i available to household k at time t; W_{jkt} is a vector of member-specific exogenous time-varying variables likely to affect nutrient requirements; X_{kt} is a vector of household- or community- specific exogenous time-varying variables; α, β and γ are parameter vectors; ε_{ijk} is the time-invariant member-specific error; μ_{iat} is the average requirements of nutrient i at time t for individuals in gender-age group a; μ_{ist} is the average requirements of nutrient i at time t for other household members with household size and demographic composition s; μ_{irt} is time-varying regional characteristics at time t affecting the intake of nutrient i for individuals in location r; and τ_{ijkt} is the remaining error.

Differencing across years within the same individual eliminates the error term ε_{ijk} in Equation (8), which could reflect unobserved activity levels or health status of each household member that persist over time.

$$\Delta\log(N_{ijkt}) = \alpha_{im}\Delta W_{jkt} + \beta_{im}\Delta\log(y_{kt}) + \gamma_{im}\Delta X_{kt} + \Delta\mu_{iat} + \Delta\mu_{ist} + \Delta\mu_{irt} + \Delta\tau_{ijkt} \qquad (9)$$

$\Delta\mu_{iat}$ reflects changes (over time) in the average nutrient requirements for individuals in gender-age group a at time t and in gender-age group a' at time $t+1$. As proxies for $\Delta\mu_{iat}$, I use gender-age group dummies (based on age in 1991) A_a. $\Delta\mu_{ist}$ represents changes (over time) in the average nutrient requirements for other household members with household size and demographic composition s. As proxies for $\Delta\mu_{ist}$, I use log household size and household demographic composition variables S_s.[11] Finally, $\Delta\mu_{irt}$ exhibits changes (over time) in regional characteristics affecting the intake of nutrient i for individuals in location r. For instance, $\Delta\mu_{irt}$ captures labor-saving technologies in agriculture that were introduced in location r between the two sample years. As proxies for $\Delta\mu_{irt}$, I use location dummies of residence R_r.[12] Thus, the estimating equation (9) can be rewritten as follows:

$$\Delta\log(N_{ijkt}) = \alpha_{im}\Delta W_{jkt} + \beta_{im}\Delta\log(y_{kt}) + \gamma_{im}\Delta X_{kt} + \delta_{Aim}A_a + \delta_{Sim}S_s + \delta_{Rim}R_r + \Delta\tau_{ijkt} \qquad (10)$$

where δ_{Aim}, δ_{Sim}, and δ_{Rim} are additional coefficient vectors. As shown in the process of deriving Equation (10), we fully control for average nutrient requirements related to age and gender. Thus, the dependent variable, changes in log nutrient intake, need not be normalized by nutrient requirements related to age and gender. This approach imposes fewer assumptions than adjusting nutrient intakes using some measure of nutrient requirements such as Chinese Dietary Reference Intakes (Chinese Nutrition Society, 2001), because it is consistent with any normalization. Similarly, changes in per-capita household nutrient $\Delta\log(y_{kt})$ need not be adjusted for changes in household demographic composition, because of the inclusion of log household size and household composition variables S_s.[13]

In the main estimating equation (10), I use per-capita household nutrients rather than per-capita household food expenditures or per-capita household income as a measure of the availability of nutrients to households. Expenditure data are not available

[11] Household size and demographic composition s are time-invariant, because I restrict the sample to those households who did not experience changes in household size and composition between the two sample years.
[12] Location of residence r is time-invariant, because I restrict the sample to those households who did not relocate from original communities between the two sample years.
[13] If we were unable to control for household demographic composition, $\Delta\log(y_{kt})$ would be positively correlated with the proportions of young children within households, which could be correlated with the individual nutrient intake of other members.

in the CHNS. Although income data are available in the CHNS, I choose per-capita nutrient intake rather than per-capita household income, because changes in income may be a poor measure of changes in total household nutrients due to both consumption smoothing and Engel's law, which suggests that the income elasticity of household food expenditures would significantly differ for households with differing wealth.

In Equation (10), there are still two important sources of bias, omitted variables and simultaneity, which need to be addressed in estimating the impact of changes in total household food consumption on the intra-household allocation of nutrients. First, the unobserved time-varying health and activity level of member j could affect not only changes in household food resources (agricultural outputs) but also changes in j's nutrient intake. Second, changes in household food resources not only affect changes in j's nutrient intake but also are affected by changes in j's nutrient intake through the effect of nutrition on productivity.

To deal with these problems, I use instruments for log per-capita household nutrient intake. The first-stage equation is specified as follows:

$$\log(y_{ikt}) = \pi_{Wim} W_{jkt} + \pi_{Xim} X_{kt} + \pi_{Zim} Z_{kt} + \theta_{ik} + \eta_{ist} + \eta_{irt} + \upsilon_{ikt} \tag{11}$$

where $\pi_{im} = (\pi_{Wim}, \pi_{Xim}, \pi_{Zim})$ is a vector of reduced-form parameters; W_{jkt} and X_{kt} are the same sets of exogenous variables used in the main estimating equation; Z_{kt} are the excluded instruments; θ_{ik} is the time-invariant household-specific error that could reflect (unobserved) permanent income or (time-invariant) consumption habit; η_{ist} is the average requirements of nutrient i at time t for households with household size and demographic composition s; η_{irt} is (time-varying) regional characteristics affecting the consumption of nutrient i at time t for households in location r; and υ_{ikt} is the remaining error. Differencing Equation (11) across years within the same household yields

$$\Delta\log(y_{ikt}) = \pi_{Wim} \Delta W_{jkt} + \pi_{Xim} \Delta X_{kt} + \pi_{Zim} \Delta Z_{kt} + \Delta\eta_{ist} + \Delta\eta_{irt} + \Delta\upsilon_{ikt} \tag{12}$$

where the time-invariant household-specific error θ_{ik} is differenced out. Using the same strategy as in the main estimating equation, we use the same set of proxies S_s and R_r for $\Delta\eta_{ist}$ and $\Delta\eta_{irt}$, respectively

$$\Delta \log(y_{ikt}) = \pi_{Wim}\Delta W_{jkt} + \pi_{Xim}\Delta X_{kt} + \pi_{Zim}\Delta Z_{kt} + \pi_{Sim}S_s + \pi_{Rim}R_r + \Delta\upsilon_{ikt} \qquad (13)$$

2SLS on Equation (10) using Equation (13) as the first-stage equation provides an unbiased estimator, as long as $\Delta\tau_{ijkt}$ and ΔZ_{kt} are not correlated.

5.2 Instruments

Rainfall is an exogenous variable that affects household food resources through its effect on agricultural income. Thus, rainfall variation is a good candidate to serve as an instrument for household food resources. We use monthly county-level rainfall data to construct instruments that capture variation in rainfall. Specifically, monthly rainfall data for the 58 sample counties are standardized using historic monthly rainfall data for the years 1961 to 1990, and the instruments are the number of standard deviations that monthly rainfall differs from historic monthly means (negative numbers if below the monthly averages).[14]

We next address potential problems with using rainfall variation as instruments for total household food consumption. It is possible that rainfall could act as a productivity shock affecting the labor supply of individuals in agriculture, which, in turn, could influence the nutrient demand of individual household members differently. This influence (through work effort) could be immediate or sequential. For instance, rainfall could change the amount of labor required in later stages of cultivation (e.g. low rainfall ruins the harvest, reducing required harvest labor, see also Fafchamps (1993) and Skoufias (1993)). To avoid these problems, I use rainfall in the previous calendar year as

[14] Historic climate data collected from more than 250 climate stations all over China are publicly available (Two Long-Term Instrumental Climatic Data Bases of the People's Republic of China, compiled by the Chinese Academy of Sciences). The University of North Carolina (UNC) merged the CHNS counties with the climate data, using an interpolation algorithm called Inverse Distance Weighting (IDW). IDW assigns the weighted average of climate data to each county, where weights are the inverses of the distances to the county from a group of surrounding climate stations located within 300km from the target county.

instruments. Nearly all surveyed households were interviewed between September and December in both 1991 and 1993. Before September of the current year, farmers have finished harvesting all crops planted in the previous calendar year. Thus, crops in the fields at the time of the CHNS interview should not be influenced by rainfall in the previous calendar year.

Rainfall in the previous calendar year should still affect current household food consumption through storage or saving, which is necessary for identification. This assumes that inter-temporal consumption smoothing is not perfect when households experience income shocks. Jalan et al (1999) and Giles (2003) both reject the hypothesis of perfect consumption smoothing for households in rural China. This is true across income levels, but especially for poor households.

We still may be concerned that past rainfall could affect current labor supply decisions (and thus nutrient demands) through other channels. First, rainfall in the previous calendar year could be correlated with current labor supply if rainfall is serially correlated. We therefore include current-year rainfall and temperature as controls. Second, previous rainfall could affect current grain prices by affecting market availability through aggregate storage. Prices affect the marginal product of labor, so could affect current labor supply. To deal with this possibility, we control for grain prices using price data available in the CHNS community survey. Finally, there remains the possibility that there is a wealth effect on labor supply, so that past shocks affecting current wealth could be correlated with current labor decisions. For example, Kochar (1999b) finds that households increase labor supply in rural India in response to idiosyncratic income shocks. Rose (2001) finds that unexpectedly bad weather and low rainfall increase labor force participation in rural India. These studies, however, look at the ex-post response of labor supply to shocks in the same cultivation year. Nonetheless, to address this possibility as well as any other indirect effects of past rainfall on current labor supply decisions, we also control directly for work hours, which we treat as endogenous.

5.3 Estimation Results

Equation (10) is estimated using 2SLS for each of the six demographic groups. We are particularly interested in the differences across the six demographic groups in the coefficient on per-capita household nutrient consumption. Per-capita household nutrient consumption y_{ikt} is just the sum of individual nutrient intakes within households, divided by household size.[15] Since my focus is on intra-household allocations, single-person households are dropped from the sample. Also, if households experienced changes in household size and/or demographic composition between 1991 and 1993, they are excluded from the sample. With these exclusions, the sample size is reduced by fifteen percent.[16]

The following variables are used as control variables. The first three sets of variables control for the possibility that the responsiveness of individual nutrient intakes differs systematically by region, participation in farming, or rural residence. The latter three sets of variables (4 - 6) control for changes in the nutrient requirements of individuals and other family members, and response differences associated with household size. Since I restrict the sample to those households who neither moved from original communities nor experienced changes in household size and demographic composition between the two sample years, location of residence and household size and composition are all time-invariant.

(1) Seven provincial dummies (excluded category: Jiangsu);

(2) Dummy for households that engaged in farming in any year of the sample period (1991 or 1993) (excluded category: never farmed);

(3) Dummy for village residents (excluded category: non-village residents);

(4) Age-group dummies created using individual ages in 1991[17] (excluded category: those aged between 30 and 32 for the prime-age equations; those aged between 60 and 62 for the elderly equations; and those aged between 15 and 17 for the child equations);

---

[15] y (per-capita household nutrient consumption) and N (individual nutritional intake) are positively correlated by construction because $y = \sum_j N_j / J$ where J is household size. Besides the theoretical grounds mentioned above, per-capita household nutrient consumption y needs to be instrumented for this reason.

[16] The main 2SLS analyses (Table 1.3) use 2806 households.

(1) Households with complete information:	3322
(2) Households without pregnant and/or lactating women in both 1991 and 1993:	2916 (87.8%)
(3) Households with condition (2) plus with multiple persons in both 1991 and 1993:	2882 (86.8%)
(4) Households with condition (3) plus without changes in household size and composition:	2806 (84.5%)

[17] 52 gender age-group dummies are created: ages 0-2, 3-5, 6-8, ..., 81-83, and 84+ for each gender.

(5) Proportions of demographic groups within households[18] (excluded category: the proportion of males aged between 25 and 50);

(6) Log household size.

Time-varying community-level controls X_{kt} include:

(7) Monthly precipitation and temperature in the current year (January to December);

(8) Log price of the major grain (either rice, flour, or corn) in the community.[19]

The major grain crop in each community is determined from the consumption data of households within the community. The coefficient on log grain price also reveals the price elasticity of the nutritional intake for each demographic group.

As instruments, the standardized amounts of rainfall in the following months of the previous calendar year are used: February, May, July, August, September, and December. First-stage results find that rainfall in the other months did not have a significant effect on household consumption.

All analyses in this study include non-farm households for the following reasons: First, some individuals work as wage laborers on farms although their households do not farm. Their current food consumption could be affected by their past wage through saving, and their past wage could be affected by past rainfall. Second, it is common for non-farm households to farm small plots for self-consumption in their spare time. It is unlikely that such crops are reported as income in the CHNS. Because non-farm households, however, would presumably be influenced by rainfall to a lesser degree than would farm households, we later restrict the sample to farm households only.

Table 1.3 presents the OLS and 2SLS coefficient estimates for per-capita household nutrient consumption for the six demographic groups (The full 2SLS results for each of the six demographic groups are reported in Appendix 1.2). Standard errors in Table 1.3 and in the rest of the chapter are all robust to household-level clustering and heteroskedasticity of any kind. Table 1.3 and the rest of the tables in this chapter contain the F statistics for the excluded instruments in the first-stage regressions and the p-values

---

[18] The proportions of 20 demographic groups are created: ages 0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-24, 25-50, 51-59, and 60+ for each gender.

[19] Only 66% of communities have the price data of the major grains both in 1991 and 1993. I imputed the missing price data, using the mean price within the county, the mean price within the province separately for rural and urban sites, and the mean price within the whole province, where the mean price within a larger area is used only when the mean price within a smaller area is not available.

of the over-identification tests for all excluded instruments, whenever 2SLS estimates are presented. All coefficient estimates are different from zero at the one percent significance level. Responses of individual nutritional intakes to changes in total household food consumption are similar for proteins and calories for a particular demographic group. According to the OLS results, males have higher elasticities than females within each generational group (prime-age, elderly, and children), and the elasticities of children and elderly members are larger and smaller than those of prime-age adults.

The 2SLS results are somewhat different. Male prime-age adults have the most elastic nutritional intakes, both for proteins (1.212) and calories (1.123). The ranking of other demographic groups from highest to lowest elasticity is as follows: girls (0.945 for proteins and 1.093 for calories); boys (0.941 for proteins and 1.013 for calories); elderly men (0.922 for proteins and 0.880 for calories); prime-age women (0.907 for proteins and 0.867 for calories); and elderly women (0.772 for proteins and 0.663 for calories). Comparing the 2SLS and OLS results for proteins, addressing the endogeneity problem raises the coefficient estimate for male prime-age adults and lowers the coefficient estimates for other demographic groups. For calories, a similar pattern is observed, but the changes in the coefficient estimates for children are more ambiguous. Male children have a slightly lower elasticity estimate with 2SLS, and female children have a higher elasticity estimate.

One possible explanation of these changes is differences across groups in physical activity. All demographic groups other than male prime-age adults could be marginal workers whose agricultural work time fluctuates, depending on shifts in labor demand. Because farm work requires greater nutrient consumption and is complementary to (unobserved) positive productivity shocks, the OLS estimates for marginal workers are biased upward. The changes in the elasticity estimate are not uniform for children because the physical activities of children (including very young boys and girls) are less influenced by agricultural work requirements. In contrast with other adults, male prime-age adults are principal workers in agricultural fields, whose physical activities are relatively less likely to be influenced by productivity shocks.[20]

[20] When we restrict the sample to include only households who say that they are farm households in at least one of the sample years (1991 and 1993), the results (Table 1.4) are similar to those in Table 1.3.

One concern about the 2SLS results is that differences in the response of nutritional intake of different demographic groups could depend on sample selection related to differences in family composition. For example, children may be more likely than the elderly to live with prime-age adults, so may have lower elasticities if prime-age adults tend to have high elasticities, even though children could have higher elasticities in families with both children and elderly members. The 2SLS results reflect both the differences in family composition of different demographic groups and the allocation decisions within households. This is an important set of parameters for evaluating the average effects of income and consumption growth on the nutrient intake of different demographic groups. However, we are also interested in understanding behavior within families. To do so, we control for differences in family composition by jointly estimating regressions for each combination of demographic groups using only households containing members of both groups. For each demographic group pair, we test formally whether elasticity differences between groups are statistically significant.

Tables 1.5 and 1.6 summarize the results of pairwise three-stage least square (3SLS) estimates for proteins and calories, respectively. The 3SLS procedure enables us to improve efficiency by taking account of cross equation error correlations, and is applied to a set of three equations: the nutrition-intake equations for two demographic groups (Equation 10 for the two demographic groups) plus the household-consumption equation (Equation 13).[21] All demographic pairs are estimated. Tables 1.5 and 1.6 show the coefficient estimates of the nutrient-intake elasticities with the corresponding standard errors in parentheses. For instance, Table 1.5 shows that 1.277 is the average elasticity of prime-age men's protein intake in households that contain both at least one male prime-age adult and one male elderly member. Tables 1.5 and 1.6 also present the differences in the elasticity estimates between the two groups being compared, the p-values of the Wald tests for equality of the elasticities of the two demographic groups, and the sample sizes. For prime-age men, all of the elasticity differences with other groups are statistically different from zero at the ten percent significance level, except for the difference with female children (not statistically significant). One disadvantage of pairwise comparisons

[21] To implement 3SLS, we require one observation per household for each equation. We thus calculate demographic group means when more than one individual in a household belongs to the same demographic group.

is that the sample sizes are much smaller. The 3SLS sample sizes for the grandparent-grandchild combinations, in particular, could be too small for meaningful analyses.

Of particular interest are tests of equality of the nutritional-intake responses across gender in the same generational group. For both protein and caloric allocations, the nutritional-intake elasticities for male prime-age adults (1.389 for proteins and 1.148 for calories) and male children (1.202 for proteins and 1.056 for calories) are significantly larger than their female counterparts (prime-age women: 0.893 for proteins and 0.918 for calories; female children: 0.753 for proteins and 0.577 for calories) at the ten percent significance level. The gender difference in the nutritional-intake response for the elderly is not statistically significant for proteins or calories.

To summarize the test results, mainly focusing on those with statistical significance, male prime-age adults have greater nutrient-intake elasticities than any other demographic group except for female children. Male children have higher elasticities than any other demographic group except for male prime-age adults. Female children have higher elasticities than demographic groups other than male prime-age adults and male children. Girls have relatively high elasticities as suggested by the 2SLS results, but when there are boys in the household, girls' elasticities become much smaller. Elderly members, regardless of gender, have lower elasticities than other demographic groups. Overall, these orderings are very similar to the 2SLS results.

5.4 Controlling for Work Hours

Next, we add current work hours as an additional endogenous independent variable. Data on farm work hours in the past week (the same period as the nutrition intake survey) are available in the CHNS. Changes in farm work hours between the two sample years are included in the differenced regression.

Table 1.7 presents the OLS and 2SLS coefficient estimates for per-capita household nutrient intake and farm work hours. Comparing Table 1.7 with Table 1.3, the results are qualitatively the same: prime-age men have the highest elasticity and the elderly have the lowest elasticities. Males have higher elasticities than females except for children's caloric-intake. Further, except for female children's protein intake, the coefficient estimates of the nutritional intakes in Table 1.7 are all within one standard error of the corresponding point estimates reported in Table 1.3. The elasticity of the protein intake of female children is only marginally lower than the point estimate minus one standard error, reported in Table 1.3.

The tests of excluded instruments in the first-stage regressions suggest that past rainfall does predict current farm work hours, especially for prime-age adults of both sexes. This could happen if wealth affects hours worked. However, even if past rainfall affects labor supply through a wealth effect, accounting for this effect by controlling for farm work hours does not alter our results, suggesting that the effect is not significantly different across demographic groups within households.[22] The 3SLS results that endogenously control for work hours (not reported) are also qualitatively the same as the results in Tables 1.5 and 1.6, although the results of the tests of equal elasticities are less sharp due to larger standard errors.[23]

5.5 Do Results Differ for the Poor?

We are concerned that the income responsiveness of the intra-household allocation of resources may depend on the level of food resources available to the household. For example, one might hypothesize that the poor always allocate food proportionally evenly to ensure subsistence (all elasticities converge to one), or less equitably, choosing to invest in stronger or favored members as incomes rise (prime-age men's elasticity is even higher). To test whether poor households behave differently than richer households, we divide the sample households into halves, depending on whether per-capita household nutrient intake is above or below the median of per-capita

[22] Instead of farm work hours, using aggregated work hours (including not only hours worked in agricultural fields but also those spent for wage employment, home gardening, livestock/poultry, fishing businesses, and small commercial household businesses) did not change the elasticity estimates meaningfully, although past rainfall did not predict the variation of aggregated work hours as much as it did the variation of farm work hours.

[23] When I control for labor supply (by two demographic groups being compared) as additional endogenous variables in the 3SLS, the coefficient estimates for the nutrient-intake elasticities change only slightly. The Hausman specification tests do not reject the hypotheses that the specifications without controlling for labor supply are adequately modeled in comparison with the specifications with endogenously controlling for labor supply. The p-values for the specification tests are all almost unity. Given larger standard errors in the 3SLS with control for labor supply, I report the results without controlling for labor supply in Tables 1.5 and 1.6.

household nutrient intake (about 61.6 g for protein and about 2114 kcal for calories),[24] and re-estimate separately for poor and richer households.

Tables 1.8 and 1.9 present the coefficient estimates and relevant statistics for proteins and calories, respectively. For both proteins and calories, the ordering of the nutrient-intake elasticity in poor households is similar to the ordering calculated using all sample households, except that girls now have the lowest elasticity estimates among all demographic groups in poor households. In richer households, the ordering is quite different. Children of both sexes have the highest elasticity estimates for both proteins and calories, except that boys have the second lowest elasticity estimate for caloric intake after elderly women. We now see that the relatively high elasticity estimates for girls when all sample households are used are driven by the high elasticity estimates for girls in richer households. A plausible story emerged from the results in Tables 1.8 and 1.9 could be that when food is relatively scarce, additional income (food) brought in households nourishes prime-age men and boys more than other household members. When food becomes relatively abundant, households begin to invest proportionally more in girls.

6. Demographic Differences in Nutrient Sufficiency

So far we have estimated demographic differences in the nutrient intake elasticity with respect to total household nutrient consumption. Although this provides an important set of parameters in contemplating income generation and transfer policies, demographic differences in the nutrient intake elasticity themselves tell us nothing about the welfare of each demographic group within households. In this section, we briefly look at how total household nutrient consumption affects the *levels* of nutrient intakes by different demographic groups.

When we discuss the levels of nutrient intakes by different family members within households, we must confront the concept of fair nutrient requirements. Nutrient requirements depend on many factors including age, sex, body size, and physical activity. To adjust individual nutrient intakes for differences in nutritional requirements related to

[24] I use the minimum of per-capita household nutrient intake of the two sample years (1991 and 1993) to allocate each household to the poor or richer group.

age, gender, and physical activity, I define Standardized Intakes (SIs), which are daily intakes normalized by Chinese Dietary Reference Intakes (DRIs):

$$\text{Standardized Intake (SI)} = \frac{\text{Daily Intake}}{\text{Chinese DRI}} \times 100$$

Chinese DRIs are developed by an association of China's foremost nutrition experts (Chinese Nutrition Society, 2001). As one can see in Table 1.10, which presents Chinese DRIs for calories and proteins, DRIs make fine distinctions by gender and age for children. For adults, DRIs differ by level of physical activity (light, moderate, heavy) in addition to age group and gender. Similarly, Household Standardized Intakes (HSIs) can be defined as follows:

$$\text{Household Standardized Intake (HSI)} = \frac{\sum_j \text{Daily Intake}_j}{\sum_j \text{DRI}_j} \times 100$$

where j indexes each household member and the summations are over all household members. As the part of the nutrition survey, the CHNS includes the physical activity level (very light, light, moderate, heavy, very heavy) of each respondent based on the person's work activity at the time of the interview.

Figures 1.1 and 1.2 present nonparametric estimates of the relationship between HSIs and SIs for proteins and calories for the six demographic groups. I use the pooled data for 1991 and 1993. To exclude outliers, only individuals from multiple-person households whose HSIs are more than 30% and less than 200% are used for the analysis.[25]

[25] The nonparametric analyses for protein and caloric intakes use 3468 households and 3479 households, respectively.

(1) Households with complete demographic information and activity level:	3664
(2) Households without pregnant and/or lactating women:	3568 (97.4%)
(3) Households with condition (2) plus with multiple persons:	3484 (95.1%)
(4) Households with condition (3) plus HSI>30 and HSI<200:	3468 (94.7%) for proteins 3479 (95.0%) for calories

Overall, Figures 1.1 and 1.2 suggest that differences in SI are larger across generations and smaller between genders within generations. Except for the lower and upper ends of HSI where the estimates could be less robust due to smaller sample sizes, the ordering of the levels of nutrient intakes are quite stable across demographic groups. For proteins, prime-age adults are favored, followed by the elderly and children, while for calories the elderly are favored, followed by prime-age adults and children.

Of course, potential confounding factors limit the validity of the nonparametric analysis. The meaningful identification in the nonparametric analysis is cross-sectional, so is subject to omitted variable bias associated with household heterogeneity, for example, with respect to health. If healthier prime-age men eat more than unhealthy men, a steeper slope of prime-age men's protein intake observed in Figure 1.1 could be due to a correlation between prime-age men's health and household wealth rather than the direct relationship between prime-age men's protein intake and household wealth.

7. Conclusions

This study examines how the intra-household allocation of nutrients responds to exogenous changes in household food consumption levels in China. I find that prime-age men have the highest elasticity of nutrient intake of all demographic groups when household food resources change exogenously. I also find that females have lower nutrient-intake elasticities than males and that elderly members have lower nutrient-intake elasticities than other groups.

These findings are somewhat at odds with existing literature that finds that human capital investments (education, medical care, and nutrients) are less income and price elastic for boys than for girls. An exception is Kochar's study (1999a), which finds that medical expenditures on prime-age men as the share of total expenditures are more income elastic than the same share for elderly men. To make sense of our results, we must return to theory. If we assume that prime-age men in China are the most productive and most favored demographic group, then our empirical findings suggest that we should question the assumption of much of the previous literature that high elasticities are an indicator of weaker status. This chapter shows that it is theoretically inconclusive whether a more productive or favored member has a higher or lower elasticity when household income changes. If elasticities, instead, are positively related to status, (as in the household model with members' utilities exhibiting decreasing relative risk aversion in consumption), then our results are consistent with gender bias, and bias against the elderly. A higher nutrient-intake elasticity for prime-age men occurs if the productivity and the marginal utility fall relatively slower for prime-age men than for other demographic groups as household resources increase. The theory also shows that predictions can be sensitive to assumptions about utilities and production. Thus, the ordering of the elasticities among demographic groups could change as household wealth increases. Further, food, as the most essential input for survival, could be allocated within families in a different manner than other human capital investments such as education.

If we focus on differences in the levels of nutrient consumption as in our non-parametric results, we find that children are worse off than prime-age adults and the elderly, and that there is no consistent gender bias within generational groups (prime-age men and boys are more favored in the protein allocation than prime-age women and girls, but the opposite is true in the caloric allocation). However, the level comparisons are more sensitive to omitted variable bias and errors in normalizing intakes for differences in nutrient requirements relating to demographic characteristics and activity levels.

Our results also deliver some policy implications. If food is given to households by government programs without any targeting effort, a larger share (not only in absolute amount but also in proportion) will go to male members than female members. Existing studies (such as Alderman and Gertler 1997 and Behrman 1988) give the impression that government food programs even without any targeting improve female nutrition more than male nutrition.

Table 1.1: Whose Nutrient Intake Is More Income Elastic? (The Case with Non-Linear Production Functions)

Table 1.1.a: Productivity in equilibrium is always equal for the prime-age and dependent members ($Y_{X_p} = Y_{X_d}$)

	$dR(X)/dX > 0$	$dR(X)/dX = 0$	$dR(X)/dX < 0$
$dQ(X)/dX > 0$	$\eta_d > \eta_p$	$\eta_d > \eta_p$	?
$dQ(X)/dX = 0$	$\eta_d > \eta_p$	$\eta_d = \eta_p$	$\eta_d < \eta_p$
$dQ(X)/dX < 0$	?	$\eta_d < \eta_p$	$\eta_d < \eta_p$

Table 1.1.b: Productivity in equilibrium is always larger for the prime-age member than for the dependent ($Y_{X_p} > Y_{X_d}$)

	$dR(X)/dX > 0$	$dR(X)/dX = 0$	$dR(X)/dX < 0$
$dQ(X)/dX > 0$	$\eta_d > \eta_p$	$\eta_d > \eta_p$	?
$dQ(X)/dX = 0$	$\eta_d > \eta_p$	$\eta_d > \eta_p$	?
$dQ(X)/dX < 0$	?	?	?

Table 1.1.c: Productivity in equilibrium is always larger for the dependent than for the prime-age member ($Y_{X_p} < Y_{X_d}$)

	$dR(X)/dX > 0$	$dR(X)/dX = 0$	$dR(X)/dX < 0$
$dQ(X)/dX > 0$	?	?	?
$dQ(X)/dX = 0$	?	$\eta_d < \eta_p$	$\eta_d < \eta_p$
$dQ(X)/dX < 0$	?	$\eta_d < \eta_p$	$\eta_d < \eta_p$

Table 1.2: Summary Statistics of Multiple-Person Households by Income Group
Means and standard deviations are shown.

	Income Group		
	Low	Middle	High
# households	2049-2069	2131-2138	2150-2157
Per-capita deflated hh income (yuan)	370.28	1,016.19	2,388.51
	(197.38)	(208.38)	(1,234.71)
household size	4.81	4.62	4.18
	(1.53)	(1.55)	(1.52)
% males aged 0-6 years in 1991	6.17	5.18	4.49
	(11.48)	(10.36)	(10.05)
% females aged 0-6 years in 1991	5.59	4.63	3.02
	(11.03)	(9.94)	(8.47)
% males aged 7-17 years in 1991	11.68	10.52	8.62
	(14.71)	(14.26)	(13.53)
% females aged 7-17 years in 1991	10.58	9.38	7.12
	(14.16)	(13.44)	(12.36)
% males aged 18-59 years in 1991	26.74	28.62	31.16
	(14.09)	(14.26)	(15.71)
% females aged 18-59 years in 1991	27.25	29.65	32.49
	(13.74)	(13.78)	(15.41)
% males aged 60+ in 1991	5.61	5.63	6.56
	(12.16)	(11.98)	(13.34)
% females aged 60+ in 1991	6.37	6.39	6.53
	(12.58)	(12.50)	(13.18)
max education within hh, % no primary education	18.01	10.58	7.91
max education within hh, % primary education	20.58	14.70	11.26
max education within hh, % more than primary education	61.40	74.72	80.84
% ever farmed in either 89, 91, or 93	86.47	65.95	46.34
% village residents	73.18	53.51	28.33
% from Liaoning	8.02	12.44	16.18
% from Henan	19.28	10.62	5.93
% from Shandong	13.10	12.96	12.47
% from Hubei	10.15	14.45	12.61
% from Hunan	11.60	10.66	16.27
% from Jiangsu	9.09	9.35	17.62
% from Guangxi	14.79	14.83	9.50
% from Guizhou	13.97	14.69	9.41
% from year 1991	50.46	50.37	49.70
% from year 1993	49.54	49.63	50.30

1) The pooled sample households for 1991 and 1993 are used.
2) Single-person households are excluded.
3) Sample sizes are slightly different for different variables.

Table 1.3: Elasticity Estimates of Nutritional Intakes by Demographic Group

demographic group	sample size	**Proteins**			**Calories**		
		OLS	2SLS		OLS	2SLS	
Male prime-age	2633	0.986	1.212	F(6, 2248)=6.18	0.974	1.123	F(6, 2248)=6.28
		(0.013)	(0.089)	Chi-sq(5) p-val=0.843	(0.017)	(0.105)	Chi-sq(5) p-val=0.927
Male elderly	450	0.939	0.922	F(6, 447)=4.25	0.937	0.880	F(6, 447)=3.38
		(0.024)	(0.100)	Chi-sq(5) p-val=0.105	(0.034)	(0.161)	Chi-sq(5) p-val=0.224
Male children	1728	1.040	0.941	F(6, 1410)=3.33	1.022	1.013	F(6, 1410)=3.82
		(0.019)	(0.140)	Chi-sq(5) p-val=0.498	(0.024)	(0.169)	Chi-sq(5) p-val=0.328
Female prime-age	2929	0.971	0.907	F(6, 2511)=6.41	0.957	0.867	F(6, 2511)=7.42
		(0.012)	(0.083)	Chi-sq(5) p-val=0.725	(0.016)	(0.091)	Chi-sq(5) p-val=0.777
Female elderly	496	0.930	0.772	F(6, 489)=3.18	0.959	0.663	F(6, 489)=2.18
		(0.026)	(0.141)	Chi-sq(5) p-val=0.298	(0.035)	(0.242)	Chi-sq(5) p-val=0.302
Female children	1521	0.977	0.945	F(6, 1219)=3.12	0.995	1.093	F(6, 1219)=3.01
		(0.021)	(0.154)	Chi-sq(5) p-val=0.346	(0.025)	(0.198)	Chi-sq(5) p-val=0.501

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) F statistics are for tests of excluded instruments on the first-stage regressions.
3) Chi-square statistics are for over-identification tests of all excluded instruments.

Table 1.4: Elasticity Estimates of Nutritional Intakes by Demographic Group (Farm Households Only)

demographic group	sample size	Proteins			Calories		
		OLS	2SLS		OLS	2SLS	
Male prime-age	1798	0.982	1.270	F(6, 1541)=4.75	0.970	1.214	F(6, 1541)=4.39
		(0.017)	(0.108)	Chi-sq(5) p-val=0.222	(0.023)	(0.133)	Chi-sq(5) p-val=0.227
Male elderly	228	0.935	0.882	F(6, 225)=3.06	0.927	0.773	F(6, 225)=2.32
		(0.036)	(0.133)	Chi-sq(5) p-val=0.177	(0.046)	(0.220)	Chi-sq(5) p-val=0.092
Male children	1333	1.033	0.910	F(6, 1054)=3.06	1.002	0.877	F(6, 1054)=3.32
		(0.022)	(0.146)	Chi-sq(5) p-val=0.748	(0.029)	(0.184)	Chi-sq(5) p-val=0.779
Female prime-age	1990	0.966	0.900	F(6, 1714)=5.52	0.952	0.803	F(6, 1714)=5.39
		(0.015)	(0.089)	Chi-sq(5) p-val=0.225	(0.021)	(0.108)	Chi-sq(5) p-val=0.370
Female elderly	264	0.891	0.732	F(6, 260)=1.98	0.948	0.751	F(6, 260)=1.87
		(0.039)	(0.173)	Chi-sq(5) p-val=0.822	(0.052)	(0.274)	Chi-sq(5) p-val=0.704
Female children	1168	0.989	0.939	F(6, 902)=3.17	1.009	1.112	F(6, 902)=2.04
		(0.025)	(0.151)	Chi-sq(5) p-val=0.245	(0.030)	(0.231)	Chi-sq(5) p-val=0.339

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) F statistics are for tests of excluded instruments on the first-stage regressions.
3) Chi-square statistics are for over-identification tests of all excluded instruments.

Table 1.5: Pair-wise 3SLS Results on Elasticities of Protein Intakes by Demographic Group

Group 1 Elasticity		Group 2 Elasticity		Difference	p-value	N
M prime-age	1.277 (0.111)	M elderly	0.654 (0.111)	0.623***	0.0004	193
M prime-age	1.515 (0.134)	M children	0.812 (0.152)	0.703***	0.0029	1125
M prime-age	1.389 (0.081)	F prime-age	0.893 (0.075)	0.496***	0.0000	2062
M prime-age	1.180 (0.116)	F elderly	0.470 (0.116)	0.710***	0.0001	260
M prime-age	0.890 (0.127)	F children	0.938 (0.173)	-0.048	0.8487	971
M elderly	0.561 (0.116)	M children	1.382 (0.166)	-0.820***	0.0001	104
M elderly	0.741 (0.111)	F prime-age	0.918 (0.092)	-0.177	0.2490	241
M elderly	0.904 (0.107)	F elderly	1.077 (0.100)	-0.173	0.2435	244
M elderly	0.221 (0.099)	F children	1.230 (0.150)	-1.009***	0.0000	79
M children	1.092 (0.144)	F prime-age	0.829 (0.100)	0.263	0.1768	1280
M children	1.447 (0.177)	F elderly	0.980 (0.291)	0.467	0.1818	149
M children	1.202 (0.121)	F children	0.753 (0.141)	0.449***	0.0092	611
F prime-age	0.689 (0.141)	F elderly	0.839 (0.161)	-0.150	0.5029	249
F prime-age	0.765 (0.129)	F children	1.168 (0.151)	-0.403*	0.0627	1089
F elderly	0.306 (0.159)	F children	1.086 (0.158)	-0.780***	0.0007	126

Table 1.6: Pair-wise 3SLS Results on Elasticities of Caloric Intakes by Demographic Group

Group 1 Elasticity		Group 2 Elasticity		Difference	p-value	N
M prime-age	1.266 (0.147)	M elderly	0.619 (0.157)	0.647***	0.0075	193
M prime-age	1.321 (0.148)	M children	0.848 (0.161)	0.472*	0.0644	1125
M prime-age	1.148 (0.088)	F prime-age	0.918 (0.083)	0.230*	0.0658	2062
M prime-age	1.602 (0.157)	F elderly	0.102 (0.201)	1.500***	0.0000	260
M prime-age	0.923 (0.153)	F children	1.110 (0.175)	-0.187	0.4901	971
M elderly	0.074 (0.145)	M children	1.286 (0.214)	-1.211***	0.0000	104
M elderly	0.210 (0.188)	F prime-age	1.061 (0.151)	-0.851**	0.0013	241
M elderly	1.052 (0.182)	F elderly	1.021 (0.142)	0.031	0.8999	244
M elderly	-0.028 (0.173)	F children	2.175 (0.205)	-2.203***	0.0000	79
M children	1.143 (0.156)	F prime-age	0.793 (0.116)	0.350	0.1059	1280
M children	1.150 (0.217)	F elderly	0.959 (0.281)	0.190	0.6104	149
M children	1.056 (0.164)	F children	0.577 (0.168)	0.479**	0.0261	611
F prime-age	0.478 (0.205)	F elderly	0.654 (0.219)	-0.176	0.5668	249
F prime-age	0.893 (0.143)	F children	1.112 (0.168)	-0.219	0.3689	1089
F elderly	0.160 (0.213)	F children	0.977 (0.183)	-0.817***	0.0040	126

1) Standard errors are in parentheses.

2) Statistically significant at the 10% *; 5% **; and 1% *** levels

Table 1.7: Elasticity Estimates of Nutritional Intakes by Demographic Group with Endogenous Labor Supply

Coefficients of Per-capita Nutrient Intake

demographic group	sample size	**Proteins**			**Calories**		
		OLS	2SLS		OLS	2SLS	
Male prime-age	2513	0.991	1.236	F(6, 2144)=4.91	0.979	1.157	F(6, 2144)=4.86
		(0.014)	(0.104)	Chi-sq(4) p-val=0.793	(0.018)	(0.134)	Chi-sq(4) p-val=0.810
Male elderly	442	0.934	0.829	F(6, 439)=4.02	0.918	0.773	F(6, 439)=3.16
		(0.024)	(0.140)	Chi-sq(4) p-val=0.108	(0.034)	(0.200)	Chi-sq(4) p-val=0.176
Male children	1720	1.040	0.916	F(6, 1404)=3.34	1.022	1.008	F(6, 1404)=3.83
		(0.019)	(0.153)	Chi-sq(4) p-val=0.432	(0.025)	(0.178)	Chi-sq(4) p-val=0.442
Female prime-age	2797	0.970	0.919	F(6, 2400)=5.03	0.955	0.856	F(6, 2400)=5.53
		(0.012)	(0.095)	Chi-sq(4) p-val=0.595	(0.017)	(0.108)	Chi-sq(4) p-val=0.608
Female elderly	490	0.927	0.716	F(6, 483)=3.40	0.964	0.577	F(6, 483)=2.42
		(0.026)	(0.181)	Chi-sq(4) p-val=0.547	(0.036)	(0.302)	Chi-sq(4) p-val=0.788
Female children	1506	0.978	0.784	F(6, 1209)=2.85	0.994	1.046	F(6, 1209)=2.79
		(0.021)	(0.260)	Chi-sq(4) p-val=0.442	(0.025)	(0.260)	Chi-sq(4) p-val=0.411

Coefficient of Farm Work Hours / 100

demographic group	sample size	**Proteins**			**Calories**		
		OLS	2SLS		OLS	2SLS	
Male prime-age	2513	0.015	0.145	F(6, 2144)=2.78	0.025	-0.039	F(6, 2144)=2.78
		(0.019)	(0.239)	Chi-sq(4) p-val=0.793	(0.018)	(0.217)	Chi-sq(4) p-val=0.810
Male elderly	442	-0.069	-0.463	F(6, 439)=1.63	-0.069	-0.301	F(6, 439)=1.63
		(0.068)	(0.535)	Chi-sq(4) p-val=0.108	(0.065)	(0.438)	Chi-sq(4) p-val=0.176
Male children	1720	-0.014	0.647	F(6, 1404)=1.13	-0.017	-1.109	F(6, 1404)=1.13
		(0.043)	(1.541)	Chi-sq(4) p-val=0.432	(0.043)	(1.380)	Chi-sq(4) p-val=0.442
Female prime-age	2797	0.005	-0.089	F(6, 2400)=5.35	0.004	-0.039	F(6, 2400)=5.35
		(0.018)	(0.136)	Chi-sq(4) p-val=0.595	(0.017)	(0.134)	Chi-sq(4) p-val=0.608
Female elderly	490	0.102	-1.039	F(6, 483)=1.80	0.080	-1.158	F(6, 483)=1.80
		(0.070)	(0.732)	Chi-sq(4) p-val=0.547	(0.070)	(0.751)	Chi-sq(4) p-val=0.788
Female children	1506	-0.076	0.986	F(6, 1209)=1.64	-0.083	0.198	F(6, 1209)=1.64
		(0.049)	(1.362)	Chi-sq(4) p-val=0.442	(0.051)	(1.049)	Chi-sq(4) p-val=0.411

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) F statistics are for tests of excluded instruments on the first-stage regressions.
3) Chi-square statistics are for over-identification tests of all excluded instruments.

Table 1.8: Elasticities of Protein Intake by Demographic Group with Endogenous Labor Supply (Poor and Richer Households Separately)

Coefficients of Per-Capita protein Intake

demographic	**Poor Households**				**Richer Households**			
group	N	OLS	2SLS		N	OLS	2SLS	
male prime-age	1186	1.013	1.382	F(6, 1045) =3.19	1327	0.939	1.183	F(6, 1098) =1.95
		(0.018)	(0.147)	Chi-sq(4) p-val=0.828		(0.019)	(0.192)	Chi-sq(4) p-val=0.341
male elderly	232	0.920	0.746	F(6, 230) =3.39	210	0.922	1.022	F(6, 208) =3.42
		(0.031)	(0.145)	Chi-sq(4) p-val=0.279		(0.042)	(0.248)	Chi-sq(4) p-val=0.601
male children	954	1.044	1.019	F(6, 762) = 2.76	766	1.030	1.664	F(6, 641) =0.80
		(0.023)	(0.191)	Chi-sq(4) p-val=0.649		(0.032)	(0.595)	Chi-sq(4) p-val=0.451
female prime-age	1388	0.960	0.900	F(6, 1208) =3.76	1409	0.984	0.753	F(6, 1191) =2.79
		(0.015)	(0.095)	Chi-sq(4) p-val=0.152		(0.019)	(0.187)	Chi-sq(4) p-val=0.734
female elderly	288	0.889	0.849	F(6, 284) =2.79	202	1.043	0.680	F(6, 198) =3.63
		(0.031)	(0.134)	Chi-sq(4) p-val=0.460		(0.048)	(0.309)	Chi-sq(4) p-val=0.776
female children	922	0.978	0.622	F(6, 714) =2.99	584	0.952	1.337	F(6, 494) =1.65
		(0.024)	(0.250)	Chi-sq(4) p-val=0.562		(0.044)	(0.324)	Chi-sq(4) p-val=0.089

Coefficient of Work Hours / 100

demographic	**Poor Households**				**Richer Households**			
group	N	OLS	2SLS		N	OLS	2SLS	
male prime-age	1186	0.014	-0.455	F(6, 1045) =2.11	1327	0.022	0.209	F(6, 1098) =1.89
		(0.028)	(0.360)	Chi-sq(4) p-val=0.828		(0.025)	(0.274)	Chi-sq(4) p-val=0.341
male elderly	232	0.002	1.299	F(6, 230) =1.38	210	-0.109	-0.525	F(6, 208) =2.61
		(0.081)	(0.667)	Chi-sq(4) p-val=0.279		(0.104)	(0.493)	Chi-sq(4) p-val=0.601
male children	954	0.108	4.452	F(6, 762) =1.78	766	-0.077	-1.173	F(6, 641) =0.92
		(0.061)	(2.459)	Chi-sq(4) p-val=0.649		(0.055)	(1.605)	Chi-sq(4) p-val=0.451
female prime-age	1388	0.019	0.060	F(6, 1208) =2.47	1409	-0.000	-0.406	F(6, 1191) =3.68
		(0.025)	(0.197)	Chi-sq(4) p-val=0.152		(0.024)	(0.201)	Chi-sq(4) p-val=0.734
female elderly	288	0.044	-1.312	F(6, 284) =0.90	202	0.213	-1.394	F(6, 198) =0.94
		(0.074)	(0.938)	Chi-sq(4) p-val=0.460		(0.138)	(0.908)	Chi-sq(4) p-val=0.776
female children	922	-0.118	2.983	F(6, 714) =1.09	584	0.041	-0.011	F(6, 494) =2.32
		(0.058)	(2.355)	Chi-sq(4) p-val=0.562		(0.095)	(0.630)	Chi-sq(4) p-val=0.089

1) Robust standard errors in parentheses are robust to household-level clustering as well as to heteroskedasticity.

2) F statistics are for tests of excluded instruments on the first-stage regressions.

3) Chi-square statistics are for over-identification tests of all excluded instruments.

Table 1.9: Elasticities of Caloric Intake by Demographic Group with Endogenous Labor Supply (Poor and Richer Households Separately)

Coefficients of Per-Capita Caloric Intake

demographic group	Poor Households N	OLS	2SLS		Richer Households N	OLS	2SLS	
Male prime-age	1171	0.996	1.228	F(6, 1038) =2.56	1342	0.933	1.106	F(6, 1105) =2.78
		(0.025)	(0.173)	Chi-sq(4) p-val=0.995		(0.020)	(0.217)	Chi-sq(4) p-val=0.774
male elderly	250	0.893	0.872	F(6, 248) =2.38	192	0.921	0.529	F(6, 190) =1.55
		(0.039)	(0.188)	Chi-sq(4) p-val=0.205		(0.079)	(0.332)	Chi-sq(4) p-val=0.324
male children	932	1.044	1.190	F(6, 756) = 2.04	788	0.964	0.498	F(6, 647) =2.86
		(0.030)	(0.257)	Chi-sq(4) p-val=0.340		(0.040)	(0.291)	Chi-sq(4) p-val=0.954
Female prime-age	1417	0.951	1.017	F(6, 1229) =3.03	1380	0.970	0.968	F(6, 1170) =2.18
		(0.021)	(0.140)	Chi-sq(4) p-val=0.101		(0.023)	(0.216)	Chi-sq(4) p-val=0.507
female elderly	303	0.942	0.724	F(6, 297) =1.44	187	1.086	0.319	F(6, 185) =1.07
		(0.038)	(0.231)	Chi-sq(4) p-val=0.294		(0.084)	(0.420)	Chi-sq(4) p-val=0.213
female children	896	0.966	0.674	F(6, 702) =3.74	610	1.046	1.298	F(6, 506) =1.11
		(0.028)	(0.180)	Chi-sq(4) p-val=0.318		(0.059)	(0.788)	Chi-sq(4) p-val=0.877

Coefficient of Work Hours / 100

demographic group	Poor Households N	OLS	2SLS		Richer Households N	OLS	2SLS	
Male prime-age	1171	0.037	-0.360	F(6, 1038) =2.79	1342	0.020	0.154	F(6, 1105) =1.93
		(0.033)	(0.304)	Chi-sq(4) p-val=0.995		(0.021)	(0.244)	Chi-sq(4) p-val=0.774
male elderly	250	-0.024	0.676	F(6, 248) =1.76	192	-0.031	-0.679	F(6, 190) =1.82
		(0.080)	(0.425)	Chi-sq(4) p-val=0.205		(0.091)	(0.440)	Chi-sq(4) p-val=0.324
male children	932	0.085	-2.307	F(6, 756) =0.99	788	-0.038	0.357	F(6, 647) =0.92
		(0.064)	(3.297)	Chi-sq(4) p-val=0.340		(0.050)	(1.029)	Chi-sq(4) p-val=0.954
Female prime-age	1417	0.003	0.097	F(6, 1229) =3.56	1380	0.011	-0.354	F(6, 1170) =2.20
		(0.027)	(0.201)	Chi-sq(4) p-val=0.101		(0.020)	(0.197)	Chi-sq(4) p-val=0.507
female elderly	303	-0.020	-0.588	F(6, 297) =0.85	187	0.106	-0.051	F(6, 185) =2.28
		(0.101)	(0.681)	Chi-sq(4) p-val=0.294		(0.103)	(0.445)	Chi-sq(4) p-val=0.213
female children	896	-0.194	2.165	F(6, 702) =1.87	610	0.001	0.205	F(6, 506) =3.14
		(0.077)	(1.613)	Chi-sq(4) p-val=0.318		(0.069)	(0.795)	Chi-sq(4) p-val=0.877

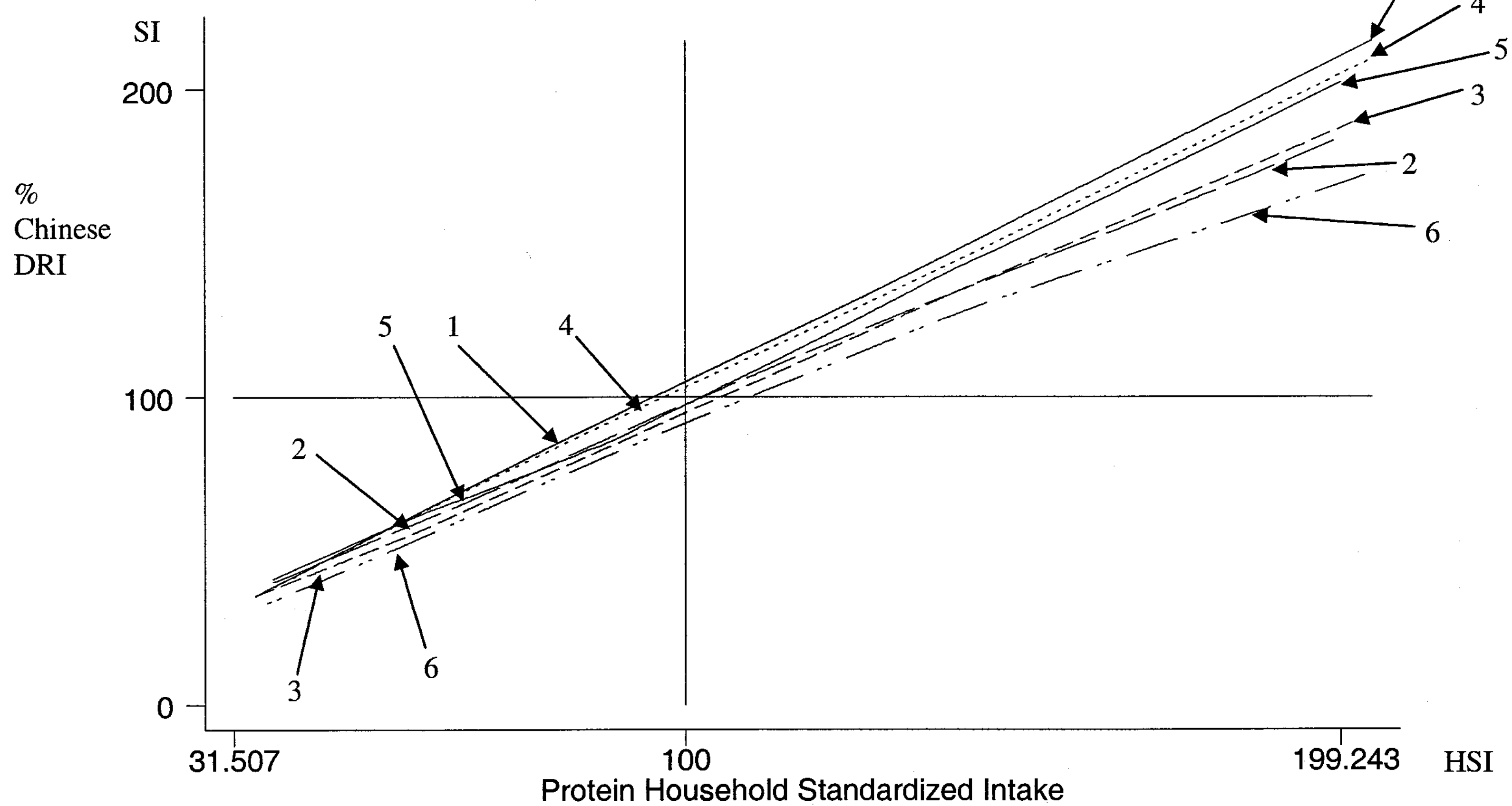
1) Robust standard errors in parentheses are robust to household-level clustering as well as to heteroskedasticity.
2) F statistics are for tests of excluded instruments on the first-stage regressions.
3) Chi-square statistics are for over-identification tests of all excluded instruments.

Table 1.10: Chinese Dietary Reference Intakes (DRIs) for Calories and Proteins

Age (year)	Calories (kcal)		Proteins (g)	
	Male	Female	Male	Female
0 ~	95kcal/kg•day	95kcal/kg•day	1.5~3g/kg•day	1.5~3g/kg•day
1 ~	1100	1050	35	35
2 ~	1200	1150	40	40
3~	1350	1300	45	45
4~	1450	1400	50	50
5~	1600	1500	55	55
6~	1700	1600	55	55
7~	1800	1700	60	60
8~	1900	1800	65	65
9~	2000	1900	65	65
10~	2100	2000	70	65
11~	2400	2200	75	75
14~	2900	2400	85	80
18~				
PAL*				
Light	2400	2100	75	65
Moderate	2700	2300	80	70
Heavy	3200	2700	90	80
Pregnant		+200		+5, +15, +20
Lactating		+500		+20
50~				
PAL*				
Light	2300	1900	75	65
Moderate	2600	2000	80	70
Heavy	3100	2200	90	80
60~			75	65
PAL*				
Light	1900	1800		
Moderate	2200	2000		
70~			75	65
PAL*				
Light	1900	1700		
Moderate	2100	1900		
80~	1900	1700	75	65

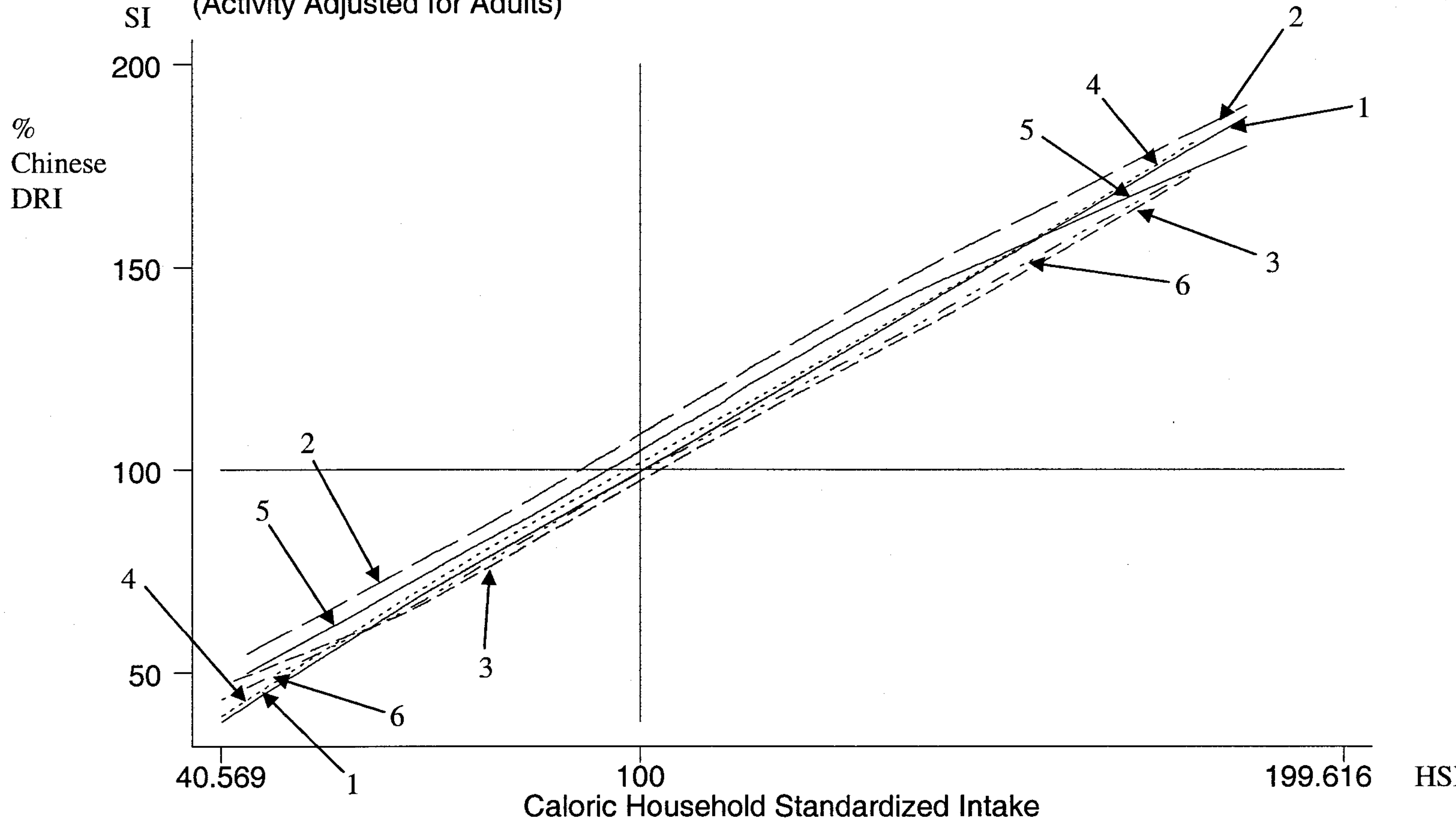
* PAL: Physical Activity Level
(Source) Chinese DRIs

Figure 1.1: Demographic Difference in Protein SI
(Activity Adjusted for Adults)



1: male prime-age, 2: male elderly, 3: male kid
4: female prime-age, 5: female elderly, 6: female kid

Figure 1.2: Demographic Difference in Caloric SI
(Activity Adjusted for Adults)



1: male prime-age, 2: male elderly, 3: male children
4: female prime-age, 5: female elderly, 6: female children

Appendix 1.1 Theory Appendix

The purpose of this appendix is three-fold. First, we show that $\eta_p = \frac{dX_p^*}{dY_0}\frac{Y_0}{X_p^*} > 0$ and $\eta_d = \frac{dX_d^*}{dY_0}\frac{Y_0}{X_d^*} > 0$ under fairly general assumptions. Second, we prove Proposition 1 used in the theory section of this chapter. Finally, we show that in the concave production case, the sign of $\eta_d - \eta_p = \frac{dX_d^*}{dY_0}\frac{Y_0}{X_d^*} - \frac{dX_p^*}{dY_0}\frac{Y_0}{X_p^*}$ is determined by the signs of both $\frac{dR(X)}{dX}$ and $\frac{dQ(X)}{dX}$ as well as other conditions, as depicted in Table 1.1.

A household solves the following problem:

$$\underset{X_p, X_d}{Max}\ U(X_p, X_d) \qquad \text{(A1)}$$

$$\text{s.t.}\quad X_p + X_d = Y(X_p, X_d, Y_0)$$

Natural assumptions are $\frac{\partial U}{\partial X_p}, \frac{\partial U}{\partial X_d} > 0$, $\frac{\partial^2 U}{\partial X_p \partial X_p}, \frac{\partial^2 U}{\partial X_d \partial X_d} < 0$, $\frac{\partial Y}{\partial X_p}, \frac{\partial Y}{\partial X_d} > 0$, and $\frac{\partial^2 Y}{\partial X_p \partial X_p}, \frac{\partial^2 Y}{\partial X_d \partial X_d} \leq 0$. Further, there is a difference in productivity between the prime-age and dependent members ($\frac{\partial Y}{\partial X_d} \leq \frac{\partial Y}{\partial X_p}$ when $X_p = X_d$). To ensure a unique solution, we also assume $0 \leq \frac{\partial Y}{\partial X_d}, \frac{\partial Y}{\partial X_p} < 1$.

The optimal X_p^* and X_d^* satisfy

$$\frac{1 - Y_{X_d}}{1 - Y_{X_p}} = \frac{U_{X_d}}{U_{X_p}} \qquad \text{(A2)}$$

$$X_p^* + X_d^* = Y(X_p^*, X_d^*, Y_0) \qquad \text{(A3)}$$

where $U_{X_p} = \frac{\partial U}{\partial X_p}$, $U_{X_d} = \frac{\partial U}{\partial X_d}$, $Y_{X_p} = \frac{\partial Y}{\partial X_p}$, and $Y_{X_d} = \frac{\partial Y}{\partial X_d}$. Taking a total derivative of (A2) with respect to Y_0 and rearranging terms yields

$$A\frac{dX_p^*}{dY_0} + B\frac{dX_d^*}{dY_0} = C \tag{A4}$$

where

$$A = -U_{X_p}Y_{X_dX_p} + (1-Y_{X_d})U_{X_pX_p} + U_{X_d}Y_{X_pX_p} - (1-Y_{X_p})U_{X_dX_p},$$

$$B = -U_{X_p}Y_{X_dX_d} + (1-Y_{X_d})U_{X_pX_d} + U_{X_d}Y_{X_pX_d} - (1-Y_{X_p})U_{X_dX_d}, \text{ and}$$

$$C = U_{X_p}Y_{X_dY_0} - U_{X_d}Y_{X_pY_0}$$

where $U_{X_dX_p} = \frac{\partial^2 U}{\partial X_d \partial X_p}$ and similarly for the other second derivatives. Taking a total derivative of (A3) with respect to Y_0 and rearranging terms yields

$$(1-Y_{X_p})\frac{dX_p^*}{dY_0} + (1-Y_{X_d})\frac{dX_d^*}{dY_0} = Y_{Y_0} \tag{A5}$$

Solving (A4) and (A5) for $\frac{dX_p^*}{dY_0}$ and $\frac{dX_d^*}{dY_0}$,

$$\frac{dX_p^*}{dY_0} = \frac{Y_{Y_0}B - (1-Y_{X_d})C}{(1-Y_{X_p})B - (1-Y_{X_d})A} \tag{A6}$$

$$\frac{dX_d^*}{dY_0} = \frac{(1-Y_{X_p})C - AY_{Y_0}}{(1-Y_{X_p})B - (1-Y_{X_d})A} \tag{A7}$$

Assuming complementarity between X_p and X_d in the household utility function ($U_{X_pX_d} \geq 0$) and in the production function ($Y_{X_pX_d} \geq 0$) as well as separability with respect to X_p and Y_0 ($Y_{X_pY_0}=0$) and with respect to X_d and Y_0 ($Y_{X_dY_0}=0$) in the production function, A, B, and C are signed unambiguously.

$$A=-U_{X_p}Y_{X_dX_p}+(1-Y_{X_d})U_{X_pX_p}+U_{X_d}Y_{X_pX_p}-(1-Y_{X_p})U_{X_dX_p}<0$$

$$B=-U_{X_p}Y_{X_dX_d}+(1-Y_{X_d})U_{X_pX_d}+U_{X_d}Y_{X_pX_d}-(1-Y_{X_p})U_{X_dX_d}>0$$

$$C=U_{X_p}Y_{X_dY_0}-U_{X_d}Y_{X_pY_0}=0$$

Then,

$$\frac{dX_p^*}{dY_0}=\frac{BY_{Y_0}}{(1-Y_{X_p})B-(1-Y_{X_d})A}>0 \tag{A8}$$

$$\frac{dX_d^*}{dY_0}=\frac{-AY_{Y_0}}{(1-Y_{X_p})B-(1-Y_{X_d})A}>0 \tag{A9}$$

These imply that the income elasticities of nutrient intakes are positive for both members.

$$\eta_m=\frac{dX_m^*}{dY_0}\frac{Y_0}{X_m^*}>0 \text{ for } m=p,d \tag{A10}$$

Next, we prove Proposition 1 used in the theory section of this chapter.[26]

Assumption 1: The household utility function is separable in X_p and X_d (i.e. $U_{X_pX_d}=0$), and individuals share a common utility function $u(X)$.

Assumption 2: The household production function is separable in X_p and X_d (i.e. $Y_{X_pX_d}=0$), and individual production functions differ only by a multiplicative constant.

Assumption 3: The individual production functions are linear (i.e. $Y_{X_pX_p}=0$ and $Y_{X_dX_d}=0$).

---

[26] Besides assumptions formally stated below, we continue to assume that $Y_{X_pY_0}=0$, $Y_{X_dY_0}=0$, and $Y_{Y_0}=1$.

It is easy to show that if the prime-age member is more productive ($\omega_p > \omega_d$) and/or more favored ($\beta < 1$), then $X_p^* > X_d^*$.

Definition: Define $R(X) = -\dfrac{u''(X)X}{u'(X)}$ *and* $Q(X) = -\dfrac{y''(X)X}{y'(X)}$ *where* $u(X)$ *and* $y(X)$ *are common individual utility and production functions defined under Assumptions 1 and 2. (For production functions,* $y(X)$ *is the common part of individual production functions without a constant.)*

Proposition 1: Under Assumptions 1, 2, 3, and $X_p^* > X_d^*$ *for any equilibrium,* $\dfrac{dR(X)}{dX} > 0$ *for any* X *implies* $\eta_d > \eta_p$, $\dfrac{dR(X)}{dX} = 0$ *for any* X *implies* $\eta_d = \eta_p$, *and* $\dfrac{dR(X)}{dX} < 0$ *for any* X *implies* $\eta_d < \eta_p$.

(Proof) Under Assumptions 1 and 2, we can calculate the difference in the income elasticity of nutrient intake for the two members, using (A8) and (A9):

$$\begin{aligned}\eta_d - \eta_p &= \frac{dX_d^*}{dY_0}\frac{Y_0}{X_d^*} - \frac{dX_p^*}{dY_0}\frac{Y_0}{X_p^*} \\ &= \frac{-Y_0}{X_p^* X_d^*}\frac{D}{(1-Y_{X_p})B - (1-Y_{X_d})A}\end{aligned} \tag{A11}$$

where

$$\begin{aligned}D &= \{X_p^*(1-Y_{X_d})U_{X_pX_p} - X_d^*(1-Y_{X_p})U_{X_dX_d}\} + \{X_p^*U_{X_d}Y_{X_pX_p} - X_d^*U_{X_p}Y_{X_dX_d}\} \\ &= (1-Y_{X_d})\{X_p^*U_{X_pX_p} - X_d^*\frac{1-Y_{X_p}}{1-Y_{X_d}}U_{X_dX_d}\} + U_{X_d}\{X_p^*Y_{X_pX_p} - X_d^*\frac{U_{X_p}}{U_{X_d}}Y_{X_dX_d}\}\end{aligned} \tag{A12}$$

The first fraction ($-Y_0 / X_p^* X_d^*$) in (A11) is unambiguously negative. The denominator in the second fraction in (A11) is unambiguously positive. D cannot be signed unambiguously, and $\eta_d - \eta_p$ and D have the opposite signs. We rewrite (A12) using (A2) as:

$$D=(1-Y_{X_d})\{X_p^*U_{X_pX_p}-X_d^*\frac{U_{X_p}}{U_{X_d}}U_{X_dX_d}\}+U_{X_d}\{X_p^*Y_{X_pX_p}-X_d^*\frac{1-Y_{X_p}}{1-Y_{X_d}}Y_{X_dX_d}\}$$

$$=(1-Y_{X_d})U_{X_p}\{\frac{X_p^*U_{X_pX_p}}{U_{X_p}}-\frac{X_d^*U_{X_dX_d}}{U_{X_d}}\}+U_{X_d}(1-Y_{X_p})\{\frac{X_p^*Y_{X_pX_p}}{1-Y_{X_p}}-\frac{X_d^*Y_{X_dX_d}}{1-Y_{X_d}}\}$$

$$=(1-Y_{X_d})U_{X_p}\{\frac{X_p^*U_{X_pX_p}}{U_{X_p}}-\frac{X_d^*U_{X_dX_d}}{U_{X_d}}\}+U_{X_d}(1-Y_{X_p})\{\frac{X_p^*Y_{X_pX_p}}{Y_{X_p}}\frac{Y_{X_p}}{1-Y_{X_p}}-\frac{X_d^*Y_{X_dX_d}}{Y_{X_d}}\frac{Y_{X_d}}{1-Y_{X_d}}\}$$

$$=-(1-Y_{X_d})U_{X_p}(R_p-R_d)-U_{X_d}(1-Y_{X_p})\{Q_p\frac{Y_{X_p}}{1-Y_{X_p}}-Q_d\frac{Y_{X_d}}{1-Y_{X_d}}\} \qquad \text{(A13)}$$

where $R_m=R(X)|_{X=X_m^*}$ and $Q_m=Q(X)|_{X=X_m^*}$ for $m=p,d$.

By Assumption 3, $Q(X)=0$, thus the last term in (A13) disappears. Under Assumptions 1, 2, 3, and $X_p^*>X_d^*$ for any equilibrium, $\frac{dR(X)}{dX}>0$ for any X implies $R_p>R_d$ for any (X_p^*,X_d^*), which is equivalent to $\eta_d>\eta_p$ for any (X_p^*,X_d^*), using (A11) and (A13). Similarly for $\frac{dR(X)}{dX}=0$ for any X and $\frac{dR(X)}{dX}<0$ for any X.

Next, we show under Assumptions 1, 2, and $X_p^*>X_d^*$ for any equilibrium that the sign of $\eta_d-\eta_p$ is determined by the signs of both $\frac{dR(X)}{dX}$ and $\frac{dQ(X)}{dX}$ as well as other conditions, as depicted in Table 1.1.[27] With a concave production function, the last term in (A13) is generally not equal to zero. Only when the two terms in (A13) agree in sign (including the cases in which one or both terms become zero), we have unambiguous predictions of the sign of $\eta_d-\eta_p$. We consider three cases. First, productivity in equilibrium is always equal for the prime-age and dependent members ($Y_{X_p}=Y_{X_d}$). Second, productivity in equilibrium is always greater for the prime-age member than for the dependent ($Y_{X_p}>Y_{X_d}$). Finally, productivity in equilibrium is always greater for the dependent than for the prime-age member ($Y_{X_p}<Y_{X_d}$).

[27] Table 1.1 applies only for the case with non-linear production functions. For the case with linear production functions, Proposition 1 applies.

In the first case ($Y_{X_p} = Y_{X_d}$), $\frac{Y_{X_p}}{1-Y_{X_p}} = \frac{Y_{X_d}}{1-Y_{X_d}}$ holds. Under Assumptions 1, 2, and $X_p^* > X_d^*$ for any equilibrium, $\frac{dQ(X)}{dX} > 0$ for any X implies $Q_p > Q_d$ for any (X_p^*, X_d^*), which is equivalent to $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} > Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$ for any (X_p^*, X_d^*). Similarly, $\frac{dQ(X)}{dX} = 0$ for any X implies $Q_p = Q_d$ for any (X_p^*, X_d^*), which is equivalent to $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} = Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$ for any (X_p^*, X_d^*), and $\frac{dQ(X)}{dX} < 0$ for any X implies $Q_p < Q_d$ for any (X_p^*, X_d^*), which is equivalent to $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} < Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$ for any (X_p^*, X_d^*). These together with a similar logic linking the sign of $\frac{dR(X)}{dX}$ to the sign of $R_p - R_d$ prove the results in Table 1.1.a.

In the second case ($Y_{X_p} > Y_{X_d}$), $\frac{Y_{X_p}}{1-Y_{X_p}} > \frac{Y_{X_d}}{1-Y_{X_d}}$ holds. Under Assumptions 1, 2, and $X_p^* > X_d^*$ for any equilibrium, $\frac{dQ(X)}{dX} > 0$ for any X implies $Q_p > Q_d$ for any (X_p^*, X_d^*), which, in turn, implies $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} > Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$ for any (X_p^*, X_d^*). Similarly, $\frac{dQ(X)}{dX} = 0$ for any X implies $Q_p = Q_d$ for any (X_p^*, X_d^*), which, in turn, implies $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} > Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$ for any (X_p^*, X_d^*). However, $\frac{dQ(X)}{dX} < 0$ does not lead to an unambiguous prediction. To see this, $\frac{dQ(X)}{dX} < 0$ for any X implies $Q_p < Q_d$ for any

(X_p^*, X_d^*), which cannot uniquely determine the sign of $Q_p \frac{Y_{X_p}}{1-Y_{X_p}} - Q_d \frac{Y_{X_d}}{1-Y_{X_d}}$. These together with a similar logic linking the sign of $\frac{dR(X)}{dX}$ to the sign of $R_p - R_d$ prove the results in Table 1.1.b. The third case ($Y_{X_p} < Y_{X_d}$) is very similar to the second case.

Finally, we briefly discuss why the measure of relative risk aversion for the production function $Q(X)$ does not produce simple conditions for the relative magnitude of η_d and η_p. As one can see in the optimality condition (A2), what matters for equilibrium is the shadow price and the marginal utility of nutrient intake. We defined $R(X)$ in terms of the marginal utility of nutrient intake. However, we did not define $Q(X)$ in terms of the shadow price of nutrient intake but defined it in terms of productivity. This complicates Table 1.1.

Define $S_m(X_m)$ that measures the analogous concepts to necessities and luxuries for the shadow price of nutrient intake by household member m.

$$S_m(X_m) = \frac{d(1-Y_{X_m})}{dX_m}\frac{X_m}{1-Y_{X_m}} = \frac{-Y_{X_m X_m} X_m}{1-Y_{X_m}} \text{ for } m = p,d \tag{A14}$$

$$S_m = \left.\frac{-Y_{X_m X_m} X_m}{1-Y_{X_m}}\right|_{X_m = X_m^*} \text{ for } m = p,d \tag{A15}$$

Because the shadow price of nutrient intake is an increasing function of nutrient intake, there is no need to put a minus sign at the front. Unfortunately, $S(X)$ cannot be defined because $S_p(X_p)$ and $S_d(X_d)$ do not have an identical functional form under Assumption 2. The relationship between S_m and Q_m can be expressed as $S_m = Q_m \left.\frac{Y_{X_m}}{1-Y_{X_m}}\right|_{X_m = X_m^*}$ for

$m = p,d$, using (A15) and the definition of Q_m. This immediately shows that the signs of $R_p - R_d$ and $S_p - S_d$ completely determine the sign of (A13) and thus the sign of $\eta_d - \eta_p$.

Appendix 1.2 Full Results of 2SLS Regressions

Table 1.11: 2SLS Regressions of Individual Protein Intakes (Full Results)
(A) Main-Equation Results

	(1)	(2)	(3)	(4)	(5)	(6)
demographic group	M prime-age	M elderly	M children	F prime-age	F elderly	F children
Δlog per-capita hh protein consump.	1.212***	0.922***	0.941***	0.907***	0.772***	0.945***
	(0.089)	(0.100)	(0.140)	(0.083)	(0.141)	(0.154)
male 0-2 yrs old (dummy)			0.182***			
			(0.057)			
male 3-5 yrs old (dummy)			0.147***			
			(0.038)			
male 6-8 yrs old (dummy)			0.090***			
			(0.032)			
male 9-11 yrs old (dummy)			0.054*			
			(0.030)			
male 12-14 yrs old (dummy)			0.061***			
			(0.023)			
male 18-20 yrs old (dummy)	-0.004					
	(0.026)					
male 21-23 yrs old (dummy)	0.013					
	(0.027)					
male 24-26 yrs old (dummy)	0.051**					
	(0.024)					
male 27-29 yrs old (dummy)	0.020					
	(0.020)					
male 33-35 yrs old (dummy)	-0.003					
	(0.020)					
male 36-38 yrs old (dummy)	-0.008					
	(0.022)					
male 39-41 yrs old (dummy)	-0.012					
	(0.022)					
male 42-44 yrs old (dummy)	-0.033					
	(0.023)					
male 45-47 yrs old (dummy)	-0.027					
	(0.028)					
male 48-50 yrs old (dummy)	-0.003					
	(0.026)					
male 51-53 yrs old (dummy)	0.029					
	(0.031)					
male 54-56 yrs old (dummy)	0.016					
	(0.031)					
male 57-59 yrs old (dummy)	0.007					
	(0.030)					
male 63-65 yrs old (dummy)		-0.017				
		(0.027)				
male 66-68 yrs old (dummy)		-0.025				
		(0.034)				
male 69-71 yrs old (dummy)		-0.052				
		(0.035)				
male 72-74 yrs old (dummy)		-0.029				
		(0.035)				

Table 1.11 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
male 75-77 yrs old (dummy)		-0.074*				
		(0.041)				
male 78-80 yrs old (dummy)		0.040				
		(0.057)				
male 81-83 yrs old (dummy)		-0.144				
		(0.113)				
male 84+ yrs old (dummy)		0.133				
		(0.113)				
female 0-2 yrs old (dummy)						0.158***
						(0.052)
female 3-5 yrs old (dummy)						0.150***
						(0.036)
female 6-8 yrs old (dummy)						0.077**
						(0.035)
female 9-11 yrs old (dummy)						0.050*
						(0.029)
female 12-14 yrs old (dummy)						0.055**
						(0.025)
female 18-20 yrs old (dummy)				0.009		
				(0.022)		
female 21-23 yrs old (dummy)				0.008		
				(0.024)		
female 24-26 yrs old (dummy)				-0.016		
				(0.020)		
female 27-29 yrs old (dummy)				-0.007		
				(0.019)		
female 33-35 yrs old (dummy)				0.014		
				(0.017)		
female 36-38 yrs old (dummy)				0.009		
				(0.019)		
female 39-41 yrs old (dummy)				0.005		
				(0.021)		
female 42-44 yrs old (dummy)				0.014		
				(0.022)		
female 45-47 yrs old (dummy)				-0.005		
				(0.024)		
female 48-50 yrs old (dummy)				0.013		
				(0.023)		
female 51-53 yrs old (dummy)				-0.013		
				(0.025)		
female 54-56 yrs old (dummy)				-0.029		
				(0.026)		
female 57-59 yrs old (dummy)				-0.016		
				(0.027)		
female 63-65 yrs old (dummy)					0.024	
					(0.028)	

Table 1.11 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
female 66-68 yrs old (dummy)					-0.017	
					(0.029)	
female 69-71 yrs old (dummy)					-0.034	
					(0.032)	
female 72-74 yrs old (dummy)					-0.027	
					(0.047)	
female 75-77 yrs old (dummy)					-0.050	
					(0.045)	
female 78-80 yrs old (dummy)					0.043	
					(0.057)	
female 81-83 yrs old (dummy)					0.111	
					(0.072)	
female 84+ yrs old (dummy)					-0.045	
					(0.090)	
ever farm (dummy)	-0.010	-0.024	0.026	0.003	-0.020	-0.007
	(0.012)	(0.033)	(0.020)	(0.011)	(0.028)	(0.021)
village resident (dummy)	0.017	0.043	0.003	-0.021**	0.031	0.032
	(0.013)	(0.034)	(0.021)	(0.011)	(0.031)	(0.021)
Δcurrent rain-January	0.090***	-0.054	-0.079*	0.011	-0.036	-0.134**
	(0.032)	(0.075)	(0.045)	(0.029)	(0.074)	(0.056)
Δcurrent rain-February	0.051*	-0.043	-0.023	-0.031	-0.154*	0.051
	(0.029)	(0.065)	(0.047)	(0.025)	(0.084)	(0.050)
Δcurrent rain-March	-0.012	-0.001	0.030**	-0.009	0.031	0.021
	(0.010)	(0.023)	(0.014)	(0.008)	(0.021)	(0.017)
Δcurrent rain-April	-0.022	-0.025	0.051	0.001	-0.008	0.006
	(0.028)	(0.058)	(0.050)	(0.026)	(0.068)	(0.047)
Δcurrent rain-May	-0.024	-0.171**	0.073	-0.024	0.030	-0.029
	(0.033)	(0.082)	(0.050)	(0.028)	(0.089)	(0.057)
Δcurrent rain-June	0.037**	0.057	-0.059**	0.012	-0.010	-0.075*
	(0.019)	(0.046)	(0.028)	(0.017)	(0.040)	(0.038)
Δcurrent rain-July	-0.006	-0.014	0.015	0.004	0.051	-0.034
	(0.022)	(0.060)	(0.030)	(0.017)	(0.066)	(0.038)
Δcurrent rain-August	0.027	-0.096	-0.042	0.017	-0.009	-0.165**
	(0.044)	(0.082)	(0.077)	(0.041)	(0.101)	(0.075)
Δcurrent rain-September	0.017	-0.004	-0.045	0.009	0.068	-0.129***
	(0.026)	(0.064)	(0.037)	(0.022)	(0.059)	(0.042)
Δcurrent rain-October	-0.031	0.033	-0.029	0.021	-0.067	0.062
	(0.032)	(0.065)	(0.062)	(0.030)	(0.071)	(0.056)
Δcurrent rain-November	-0.004	-0.057	0.083**	0.009	-0.013	0.015
	(0.020)	(0.043)	(0.035)	(0.019)	(0.049)	(0.032)
Δcurrent rain-December	0.044	-0.210**	0.050	-0.001	-0.210*	0.037
	(0.040)	(0.103)	(0.061)	(0.035)	(0.110)	(0.075)
Δcurrent temperature-Jan	0.080	0.142	0.101	0.024	-0.240	-0.029
	(0.063)	(0.159)	(0.105)	(0.056)	(0.163)	(0.116)
Δcurrent temperature-Feb	0.095	-0.618**	0.317*	0.024	-0.452	-0.110
	(0.132)	(0.242)	(0.175)	(0.111)	(0.312)	(0.181)

Table 1.11 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Δcurrent temperature-Mar	-0.168*	-0.139	0.358*	-0.128	0.266	0.370**
	(0.099)	(0.218)	(0.183)	(0.093)	(0.218)	(0.181)
Δcurrent temperature-Apr	0.037	-0.058	0.102	-0.071	-0.109	0.042
	(0.055)	(0.104)	(0.096)	(0.047)	(0.099)	(0.092)
Δcurrent temperature-May	0.036	-0.103	-0.026	-0.010	-0.071	-0.051
	(0.074)	(0.156)	(0.112)	(0.071)	(0.191)	(0.130)
Δcurrent temperature-Jun	0.048	-0.041	0.085	0.008	-0.178*	-0.019
	(0.045)	(0.084)	(0.068)	(0.039)	(0.093)	(0.070)
Δcurrent temperature-Jul	-0.079**	0.015	0.021	0.011	0.180*	0.020
	(0.037)	(0.090)	(0.053)	(0.032)	(0.098)	(0.064)
Δcurrent temperature-Aug	-0.041	-0.090	0.101	0.003	-0.041	0.047
	(0.044)	(0.117)	(0.063)	(0.039)	(0.102)	(0.084)
Δcurrent temperature-Sep	0.149	-0.070	0.039	0.065	-0.298	-0.302*
	(0.099)	(0.210)	(0.142)	(0.099)	(0.250)	(0.156)
Δcurrent temperature-Oct	0.042	0.234	-0.268	0.028	0.110	-0.238
	(0.106)	(0.189)	(0.193)	(0.094)	(0.199)	(0.173)
Δcurrent temperature-Nov	0.004	-0.203	0.117	-0.065	0.051	0.107
	(0.074)	(0.143)	(0.115)	(0.064)	(0.146)	(0.125)
Δcurrent temperature-Dec	0.073	-0.119	-0.104	-0.007	0.175	-0.065
	(0.062)	(0.139)	(0.109)	(0.052)	(0.134)	(0.127)
Δlog price of most eaten grain in	0.031*	0.054	-0.001	0.010	-0.107**	-0.062*
community (rice, flour, or corn)	(0.017)	(0.045)	(0.028)	(0.017)	(0.051)	(0.032)
share males 0-2 yrs old	0.027	0.210	-0.178	-0.127	0.123	-0.078
	(0.095)	(0.353)	(0.213)	(0.078)	(0.245)	(0.204)
share males 3-5 yrs old	-0.114	0.157	-0.142	-0.219***	-0.458**	0.018
	(0.073)	(0.151)	(0.156)	(0.071)	(0.188)	(0.200)
share males 6-8 yrs old	-0.076	-0.073	-0.176	-0.109*	0.018	-0.182
	(0.072)	(0.170)	(0.150)	(0.062)	(0.201)	(0.175)
share males 9-11 yrs old	-0.026	-0.148	-0.074	-0.188***	-0.177	0.038
	(0.073)	(0.162)	(0.152)	(0.065)	(0.164)	(0.169)
share males 12-14 yrs old	-0.032	-0.081	-0.146	-0.120*	-0.011	0.010
	(0.068)	(0.271)	(0.138)	(0.061)	(0.255)	(0.172)
share males 15-17 yrs old	0.021	0.013	-0.013	-0.108*	0.103	0.032
	(0.064)	(0.199)	(0.127)	(0.055)	(0.223)	(0.183)
share males 18-24 yrs old	0.027	0.049	0.117	-0.006	-0.450***	0.036
	(0.054)	(0.110)	(0.137)	(0.047)	(0.114)	(0.161)
share males 51-59 yrs old	-0.050	-0.555	0.006	-0.004	0.057	0.024
	(0.077)	(0.348)	(0.152)	(0.055)	(0.150)	(0.139)
share males 60+ yrs old	-0.063	-0.201	-0.084	-0.032	-0.037	0.298**
	(0.077)	(0.389)	(0.144)	(0.057)	(0.102)	(0.151)
share females 0-2 yrs old	0.033	0.028	0.229	-0.033	-0.075	0.159
	(0.093)	(0.267)	(0.169)	(0.079)	(0.198)	(0.215)
share females 3-5 yrs old	-0.035	-0.155	-0.032	-0.148**	-0.113	-0.061
	(0.070)	(0.346)	(0.145)	(0.062)	(0.191)	(0.178)
share females 6-8 yrs old	-0.012	-0.062	-0.171	-0.091	-0.254	0.037
	(0.073)	(0.303)	(0.155)	(0.068)	(0.242)	(0.172)

Table 1.11 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
share females 9-11 yrs old	-0.116	-0.072	0.057	-0.153**	0.058	0.023
	(0.076)	(0.207)	(0.136)	(0.061)	(0.264)	(0.169)
share females 12-14 yrs old	-0.013	0.372	-0.117	-0.133**	0.333	-0.120
	(0.071)	(0.320)	(0.135)	(0.060)	(0.234)	(0.159)
share females 15-17 yrs old	0.073	-0.051	-0.080	-0.045	-0.095	-0.019
	(0.070)	(0.186)	(0.131)	(0.054)	(0.213)	(0.157)
share females 18-24 yrs old	-0.006	0.045	-0.127	-0.067	-0.118	-0.056
	(0.066)	(0.134)	(0.129)	(0.054)	(0.124)	(0.171)
share females 25-50 yrs old	-0.022	0.003	-0.014	-0.046	-0.116	0.202
	(0.080)	(0.114)	(0.199)	(0.067)	(0.131)	(0.239)
share females 51-59 yrs old	0.015	0.096	0.214	-0.024	0.373	0.149
	(0.086)	(0.166)	(0.176)	(0.080)	(0.342)	(0.253)
share females 60+ yrs old	-0.052	0.136	0.153	-0.171**	-0.182	0.090
	(0.077)	(0.160)	(0.153)	(0.069)	(0.246)	(0.167)
log household size	-0.007	-0.040	0.022	0.022	-0.034	0.022
(no change in hh size b/w 91 & 93)	(0.016)	(0.111)	(0.035)	(0.015)	(0.073)	(0.035)
Liaoning province (dummy)	0.005	-0.014	-0.177	-0.045	0.019	-0.130
	(0.156)	(0.408)	(0.249)	(0.144)	(0.420)	(0.240)
Henan province (dummy)	0.113	-0.130	-0.030	0.099	-0.312	-0.272*
	(0.091)	(0.215)	(0.137)	(0.079)	(0.236)	(0.155)
Shandong province (dummy)	0.063	-0.013	-0.322***	0.077	-0.009	-0.335***
	(0.069)	(0.173)	(0.107)	(0.063)	(0.190)	(0.116)
Hubei province (dummy)	0.059	-0.375	0.066	0.073	-0.226	-0.261
	(0.108)	(0.254)	(0.152)	(0.092)	(0.261)	(0.190)
Hunan province (dummy)	0.216	-0.585*	0.029	0.099	-0.202	-0.345
	(0.140)	(0.307)	(0.198)	(0.119)	(0.316)	(0.237)
Guangxi province (dummy)	0.642**	-0.443	0.222	0.001	-1.103**	-0.494
	(0.263)	(0.546)	(0.354)	(0.233)	(0.543)	(0.389)
Guizhou province (dummy)	0.160	-0.603*	0.416	0.069	0.105	-0.238
	(0.180)	(0.362)	(0.268)	(0.167)	(0.402)	(0.313)
Constant	-0.103	0.830	-0.394	0.112	0.489	-0.074
	(0.332)	(0.608)	(0.552)	(0.274)	(0.626)	(0.443)
Sample size	2633	450	1728	2929	496	1521
p-value of Hansen J statistic (over-identification test statistic)	0.84	0.10	0.50	0.73	0.30	0.35

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.12: 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
F-statistic on the excluded IV's	F(6,2248) = 6.18	F(6, 447) = 4.25	F(6,1410) = 3.33	F(6,2511) = 6.41	F(6,489) = 3.18	F(6,1219) = 3.12
p-value joint significance	0.0000	0.0004	0.0029	0.0000	0.0046	0.0049
Δprevious rain-February	-0.262***	-0.715***	-0.177**	-0.257***	-0.558***	-0.158
	(0.064)	(0.154)	(0.085)	(0.062)	(0.153)	(0.099)
Δprevious rain-May	-0.414***	-0.748***	-0.367***	-0.426***	-0.507**	-0.353***
	(0.083)	(0.198)	(0.110)	(0.082)	(0.203)	(0.110)
Δprevious rain-July	0.203***	0.103	0.243**	0.252***	0.239*	0.358***
	(0.070)	(0.139)	(0.096)	(0.069)	(0.138)	(0.098)
Δprevious rain-August	0.261**	0.218	0.329**	0.251**	0.290	0.289*
	(0.103)	(0.250)	(0.148)	(0.100)	(0.253)	(0.152)
Δprevious rain-September	0.132**	0.120	0.134**	0.129**	0.187	0.156**
	(0.053)	(0.123)	(0.065)	(0.052)	(0.128)	(0.074)
Δprevious rain-December	0.501***	1.004***	0.376***	0.481***	0.608**	0.383***
	(0.104)	(0.261)	(0.143)	(0.097)	(0.249)	(0.145)
male 0-2 yrs old (d.)			-0.062			
			(0.054)			
male 3-5 yrs old (d.)			-0.009			
			(0.044)			
male 6-8 yrs old (d.)			-0.019			
			(0.040)			
male 9-11 yrs old (d.)			-0.045			
			(0.037)			
male 12-14 yrs old (d.)			-0.040			
			(0.032)			
male 18-20 yrs old (d.)	0.033					
	(0.039)					
male 21-23 yrs old (d.)	0.016					
	(0.041)					
male 24-26 yrs old (d.)	0.012					
	(0.039)					
male 27-29 yrs old (d.)	0.049					
	(0.035)					
male 33-35 yrs old (d.)	0.055*					
	(0.032)					
male 36-38 yrs old (d.)	0.022					
	(0.036)					
male 39-41 yrs old (d.)	0.049					
	(0.036)					
male 42-44 yrs old (d.)	0.068*					
	(0.039)					
male 45-47 yrs old (d.)	0.058					
	(0.041)					
male 48-50 yrs old (d.)	0.035					
	(0.040)					

1) d. denotes dummy.

Table 1.12 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
male 51-53 yrs old (dummy)	0.036					
	(0.047)					
male 54-56 yrs old (dummy)	0.065					
	(0.042)					
male 57-59 yrs old (dummy)	0.072*					
	(0.039)					
male 63-65 yrs old (dummy)		0.067				
		(0.048)				
male 66-68 yrs old (dummy)		-0.052				
		(0.057)				
male 69-71 yrs old (dummy)		-0.053				
		(0.062)				
male 72-74 yrs old (dummy)		-0.031				
		(0.061)				
male 75-77 yrs old (dummy)		-0.026				
		(0.084)				
male 78-80 yrs old (dummy)		-0.101				
		(0.092)				
male 81-83 yrs old (dummy)		-0.099				
		(0.125)				
male 84+ yrs old (dummy)		0.170				
		(0.195)				
female 0-2 yrs old (dummy)						0.020
						(0.052)
female 3-5 yrs old (dummy)						-0.050
						(0.041)
female 6-8 yrs old (dummy)						0.005
						(0.038)
female 9-11 yrs old (dummy)						0.001
						(0.036)
female 12-14 yrs old (dummy)						0.026
						(0.031)
female 18-20 yrs old (dummy)				0.066*		
				(0.035)		
female 21-23 yrs old (dummy)				0.063		
				(0.040)		
female 24-26 yrs old (dummy)				0.034		
				(0.036)		
female 27-29 yrs old (dummy)				0.041		
				(0.031)		
female 33-35 yrs old (dummy)				0.069**		
				(0.032)		
female 36-38 yrs old (dummy)				0.083**		
				(0.034)		

Table 1.12 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
female 39-41 yrs old (d.)				0.100***		
				(0.035)		
female 42-44 yrs old (d.)				0.080**		
				(0.038)		
female 45-47 yrs old (d.)				0.098**		
				(0.040)		
female 48-50 yrs old (d.)				0.047		
				(0.040)		
female 51-53 yrs old (d.)				0.029		
				(0.041)		
female 54-56 yrs old (d.)				0.064		
				(0.040)		
female 57-59 yrs old (d.)				-0.018		
				(0.040)		
female 63-65 yrs old (d.)					0.068	
					(0.051)	
female 66-68 yrs old (d.)					0.027	
					(0.054)	
female 69-71 yrs old (d.)					0.041	
					(0.058)	
female 72-74 yrs old (d.)					-0.011	
					(0.073)	
female 75-77 yrs old (d.)					0.121**	
					(0.059)	
female 78-80 yrs old (d.)					0.088	
					(0.077)	
female 81-83 yrs old (d.)					0.144	
					(0.092)	
female 84+ yrs old (d.)					0.037	
					(0.072)	
ever farm (d.)	0.025	-0.112	0.012	0.031	0.017	0.037
	(0.023)	(0.069)	(0.029)	(0.023)	(0.058)	(0.034)
village resident (d.)	-0.051**	0.110	0.006	-0.022	0.039	-0.038
	(0.025)	(0.070)	(0.032)	(0.024)	(0.059)	(0.033)
Δcurrent rain-January	-0.457***	-0.995***	-0.335***	-0.430***	-0.552**	-0.326**
	(0.091)	(0.273)	(0.119)	(0.085)	(0.240)	(0.134)
Δcurrent rain-February	-0.433***	-0.740***	-0.427***	-0.443***	-0.739***	-0.336***
	(0.071)	(0.181)	(0.109)	(0.074)	(0.175)	(0.109)
Δcurrent rain-March	-0.012	-0.105*	-0.023	-0.018	-0.019	0.014
	(0.026)	(0.058)	(0.036)	(0.025)	(0.055)	(0.038)
Δcurrent rain-April	-0.067	0.092	-0.141**	-0.094*	-0.198*	-0.082
	(0.051)	(0.116)	(0.066)	(0.048)	(0.120)	(0.073)
Δcurrent rain-May	0.348***	0.794***	0.279*	0.397***	0.954***	0.409**
	(0.117)	(0.252)	(0.156)	(0.118)	(0.269)	(0.170)

1) d. denotes dummy.

Table 1.12 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Δcurrent rain-June	-0.096**	-0.225**	-0.045	-0.071*	-0.034	-0.102
	(0.044)	(0.105)	(0.056)	(0.041)	(0.102)	(0.065)
Δcurrent rain-July	0.126**	0.128	0.145**	0.139***	0.390***	0.125
	(0.049)	(0.106)	(0.063)	(0.054)	(0.110)	(0.078)
Δcurrent rain-August	-0.375***	-0.530***	-0.400***	-0.347***	-0.124	-0.254**
	(0.074)	(0.191)	(0.100)	(0.073)	(0.205)	(0.106)
Δcurrent rain-September	0.090	0.118	0.073	0.079	0.268***	0.125
	(0.058)	(0.117)	(0.070)	(0.057)	(0.098)	(0.077)
Δcurrent rain-October	0.316***	0.249	0.420***	0.323***	0.126	0.336***
	(0.083)	(0.200)	(0.114)	(0.080)	(0.225)	(0.116)
Δcurrent rain-November	-0.215***	-0.545***	-0.188**	-0.242***	-0.499***	-0.117
	(0.067)	(0.152)	(0.092)	(0.063)	(0.157)	(0.102)
Δcurrent rain-December	0.053	-0.295	0.034	-0.019	-0.465**	0.124
	(0.099)	(0.227)	(0.133)	(0.093)	(0.234)	(0.137)
Δcurrent temperature-Jan	-0.616***	-0.995**	-0.372	-0.710***	-1.431***	-0.645**
	(0.228)	(0.434)	(0.309)	(0.226)	(0.449)	(0.325)
Δcurrent temperature-Feb	1.192***	2.687**	1.332**	1.264***	2.224**	1.149*
	(0.438)	(1.078)	(0.592)	(0.429)	(1.016)	(0.601)
Δcurrent temperature-Mar	-0.788***	-1.332***	-0.795***	-0.771***	-0.652	-0.658**
	(0.248)	(0.494)	(0.293)	(0.230)	(0.470)	(0.329)
Δcurrent temperature-Apr	-0.711***	-1.222***	-0.641***	-0.655***	-0.913***	-0.532***
	(0.110)	(0.249)	(0.134)	(0.102)	(0.250)	(0.143)
Δcurrent temperature-May	0.498***	0.460	0.609***	0.659***	1.137***	0.557**
	(0.173)	(0.344)	(0.228)	(0.163)	(0.376)	(0.248)
Δcurrent temperature-Jun	0.221**	0.627**	0.221	0.261**	0.517**	0.323**
	(0.112)	(0.272)	(0.145)	(0.106)	(0.245)	(0.164)
Δcurrent temperature-Jul	0.449***	0.684**	0.407***	0.414***	0.788***	0.281*
	(0.106)	(0.268)	(0.145)	(0.104)	(0.258)	(0.155)
Δcurrent temperature-Aug	-0.275**	-0.757***	-0.318**	-0.397***	-0.767***	-0.598***
	(0.129)	(0.259)	(0.160)	(0.126)	(0.258)	(0.175)
Δcurrent temperature-Sep	-0.305	-0.486	-0.211	-0.447	-0.459	-0.154
	(0.309)	(0.661)	(0.422)	(0.306)	(0.595)	(0.455)
Δcurrent temperature-Oct	1.255***	1.620***	1.212***	1.097***	1.265***	0.777***
	(0.220)	(0.492)	(0.264)	(0.201)	(0.463)	(0.293)
Δcurrent temperature-Nov	-0.306**	-0.287	-0.239	-0.177	0.233	-0.005
	(0.153)	(0.310)	(0.176)	(0.139)	(0.320)	(0.197)
Δcurrent temperature-Dec	-0.011	0.043	-0.012	0.033	0.029	0.258
	(0.180)	(0.390)	(0.244)	(0.167)	(0.422)	(0.252)
Δlog price of most eaten grain in community (rice, flour, or corn)	-0.008	-0.014	0.045	0.018	-0.129*	-0.053
	(0.038)	(0.080)	(0.052)	(0.036)	(0.073)	(0.051)
share males 0-2 yrs old	-0.210	-0.554	-0.003	-0.164	-0.487	0.291
	(0.159)	(0.449)	(0.264)	(0.156)	(0.351)	(0.324)
share males 3-5 yrs old	0.037	-0.230	-0.048	0.106	-0.348	0.588**
	(0.131)	(0.313)	(0.228)	(0.125)	(0.253)	(0.271)
share males 6-8 yrs old	0.076	-0.106	0.000	0.052	-0.089	0.516*
	(0.133)	(0.385)	(0.217)	(0.130)	(0.383)	(0.273)

Table 1.12 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
share males 9-11 yrs old	-0.055	-0.308	0.026	-0.045	-0.273	0.447
	(0.136)	(0.282)	(0.223)	(0.132)	(0.284)	(0.277)
share males 12-14 yrs old	-0.022	0.476	0.067	-0.038	0.018	0.203
	(0.136)	(0.369)	(0.217)	(0.132)	(0.309)	(0.256)
share males 15-17 yrs old	-0.182	0.055	-0.105	-0.119	-0.339	0.351
	(0.133)	(0.306)	(0.202)	(0.125)	(0.347)	(0.275)
share males 18-24 yrs old	-0.056	-0.109	0.042	-0.143	-0.295*	0.369
	(0.108)	(0.200)	(0.205)	(0.106)	(0.166)	(0.258)
share males 51-59 yrs old	-0.212	-0.024	-0.601**	-0.035	0.275	-0.098
	(0.131)	(0.509)	(0.254)	(0.120)	(0.272)	(0.331)
share males 60+ yrs old	-0.234	-0.044	-0.200	-0.105	-0.328**	0.137
	(0.151)	(0.565)	(0.211)	(0.121)	(0.158)	(0.272)
share females 0-2 yrs old	-0.047	-0.230	0.289	-0.077	-0.057	0.345
	(0.169)	(0.408)	(0.252)	(0.163)	(0.446)	(0.315)
share females 3-5 yrs old	-0.145	0.324	-0.146	-0.127	-0.195	0.384
	(0.135)	(0.523)	(0.209)	(0.131)	(0.394)	(0.268)
share females 6-8 yrs old	-0.098	-0.682**	0.042	-0.009	-0.244	0.225
	(0.136)	(0.335)	(0.200)	(0.131)	(0.315)	(0.259)
share females 9-11 yrs old	-0.205	-0.051	-0.124	-0.207	-0.259	0.103
	(0.136)	(0.467)	(0.201)	(0.128)	(0.414)	(0.265)
share females 12-14 yrs old	-0.051	0.517	0.019	-0.140	0.520	0.113
	(0.147)	(0.437)	(0.205)	(0.134)	(0.332)	(0.275)
share females 15-17 yrs old	-0.184	-0.504	-0.140	-0.220	-0.310	0.228
	(0.152)	(0.341)	(0.220)	(0.138)	(0.331)	(0.269)
share females 18-24 yrs old	-0.228*	-0.418*	-0.084	-0.205*	-0.392*	0.159
	(0.125)	(0.227)	(0.204)	(0.115)	(0.206)	(0.288)
share females 25-50 yrs old	-0.145	-0.067	0.111	0.034	-0.183	0.718**
	(0.163)	(0.243)	(0.304)	(0.159)	(0.231)	(0.363)
share females 51-59 yrs old	-0.161	-0.535**	-0.016	0.054	-1.029**	0.539
	(0.174)	(0.248)	(0.295)	(0.176)	(0.439)	(0.416)
share females 60+ yrs old	0.072	-0.378*	0.026	-0.041	-0.015	0.162
	(0.148)	(0.225)	(0.219)	(0.133)	(0.402)	(0.285)
log household size	-0.015	0.027	0.047	0.020	0.024	0.039
(no change in hh size b/w 91 & 93)	(0.031)	(0.149)	(0.047)	(0.030)	(0.114)	(0.060)
Liaoning province (dummy)	-1.775***	-3.827***	-1.687***	-1.888***	-2.265**	-1.598**
	(0.479)	(1.211)	(0.620)	(0.453)	(1.134)	(0.650)
Henan province (dummy)	-0.589***	-0.501	-0.351	-0.500**	-0.111	-0.345
	(0.216)	(0.548)	(0.281)	(0.206)	(0.506)	(0.312)
Shandong province (dummy)	-0.060	-0.087	0.041	-0.036	0.431	0.105
	(0.174)	(0.461)	(0.228)	(0.164)	(0.396)	(0.247)
Hubei province (dummy)	-0.486**	-0.590	-0.242	-0.380	0.167	-0.365
	(0.248)	(0.616)	(0.328)	(0.239)	(0.563)	(0.358)
Hunan province (dummy)	-0.890***	-1.278*	-0.570	-0.720***	0.002	-0.647*
	(0.276)	(0.738)	(0.356)	(0.267)	(0.673)	(0.378)

Table 1.12 (continued): 2SLS Regressions of Individual Protein Intakes (Full Results)
(B) First-Stage Results

	(1)	(2)	(3)	(4)	(5)	(6)
demographic group	M prime-age	M elderly	M children	F prime-age	F elderly	F children
Guangxi province (dummy)	-1.006	-2.006	-0.099	-0.873	-0.981	-0.021
	(0.617)	(1.459)	(0.826)	(0.604)	(1.236)	(0.887)
Guizhou province (dummy)	-0.222	0.297	0.023	-0.023	1.748**	0.250
	(0.361)	(0.832)	(0.458)	(0.343)	(0.811)	(0.495)
Constant	1.353**	0.423	1.301*	0.788	-0.688	0.023
	(0.554)	(1.280)	(0.787)	(0.525)	(1.235)	(0.823)
Sample size	2633	450	1728	2929	496	1521

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.13: 2SLS Regressions of Individual Caloric Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Δlog per-capita hh caloric consump	1.123***	0.880***	1.013***	0.867***	0.663***	1.093***
	(0.105)	(0.161)	(0.169)	(0.091)	(0.242)	(0.198)
male 0-2 yrs old (dummy)			0.215***			
			(0.056)			
male 3-5 yrs old (dummy)			0.150***			
			(0.037)			
male 6-8 yrs old (dummy)			0.102***			
			(0.031)			
male 9-11 yrs old (dummy)			0.060**			
			(0.029)			
male 12-14 yrs old (dummy)			0.055**			
			(0.022)			
male 18-20 yrs old (dummy)	-0.003					
	(0.024)					
male 21-23 yrs old (dummy)	0.028					
	(0.025)					
male 24-26 yrs old (dummy)	0.042**					
	(0.021)					
male 27-29 yrs old (dummy)	0.014					
	(0.018)					
male 33-35 yrs old (dummy)	0.003					
	(0.017)					
male 36-38 yrs old (dummy)	-0.008					
	(0.019)					
male 39-41 yrs old (dummy)	-0.011					
	(0.019)					
male 42-44 yrs old (dummy)	-0.028					
	(0.020)					
male 45-47 yrs old (dummy)	-0.028					
	(0.026)					
male 48-50 yrs old (dummy)	0.000					
	(0.024)					
male 51-53 yrs old (dummy)	0.028					
	(0.026)					
male 54-56 yrs old (dummy)	0.029					
	(0.028)					
male 57-59 yrs old (dummy)	-0.004					
	(0.026)					
male 63-65 yrs old (dummy)		-0.020				
		(0.025)				
male 66-68 yrs old (dummy)		0.007				
		(0.033)				
male 69-71 yrs old (dummy)		-0.050				
		(0.033)				
male 72-74 yrs old (dummy)		-0.031				
		(0.032)				

Table 1.13 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
male 75-77 yrs old (dummy)		-0.069*				
		(0.041)				
male 78-80 yrs old (dummy)		0.023				
		(0.059)				
male 81-83 yrs old (dummy)		-0.155*				
		(0.091)				
male 84+ yrs old (dummy)		0.131				
		(0.101)				
female 0-2 yrs old (dummy)						0.181***
						(0.051)
female 3-5 yrs old (dummy)						0.150***
						(0.034)
female 6-8 yrs old (dummy)						0.083**
						(0.034)
female 9-11 yrs old (dummy)						0.045
						(0.030)
female 12-14 yrs old (dummy)						0.043*
						(0.026)
female 18-20 yrs old (dummy)				0.013		
				(0.020)		
female 21-23 yrs old (dummy)				0.010		
				(0.022)		
female 24-26 yrs old (dummy)				-0.006		
				(0.018)		
female 27-29 yrs old (dummy)				-0.012		
				(0.018)		
female 33-35 yrs old (dummy)				0.010		
				(0.015)		
female 36-38 yrs old (dummy)				0.016		
				(0.017)		
female 39-41 yrs old (dummy)				0.001		
				(0.018)		
female 42-44 yrs old (dummy)				0.011		
				(0.019)		
female 45-47 yrs old (dummy)				0.004		
				(0.020)		
female 48-50 yrs old (dummy)				0.008		
				(0.021)		
female 51-53 yrs old (dummy)				-0.002		
				(0.023)		
female 54-56 yrs old (dummy)				-0.022		
				(0.024)		

Table 1.13 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
female 57-59 yrs old (dummy)				-0.012		
				(0.027)		
female 63-65 yrs old (dummy)					0.021	
					(0.027)	
female 66-68 yrs old (dummy)					-0.051*	
					(0.030)	
female 69-71 yrs old (dummy)					-0.034	
					(0.033)	
female 72-74 yrs old (dummy)					-0.004	
					(0.050)	
female 75-77 yrs old (dummy)					-0.063	
					(0.049)	
female 78-80 yrs old (dummy)					0.037	
					(0.060)	
female 81-83 yrs old (dummy)					0.099	
					(0.076)	
female 84+ yrs old (dummy)					-0.035	
					(0.091)	
ever farm (dummy)	0.001	-0.012	0.002	-0.003	-0.011	-0.005
	(0.010)	(0.030)	(0.019)	(0.010)	(0.032)	(0.022)
village resident (dummy)	0.011	0.015	0.018	-0.018*	0.019	0.025
	(0.013)	(0.029)	(0.019)	(0.010)	(0.033)	(0.021)
Δcurrent rain-January	0.062**	-0.005	-0.037	-0.005	-0.027	-0.116**
	(0.028)	(0.072)	(0.041)	(0.027)	(0.075)	(0.050)
Δcurrent rain-February	0.011	-0.034	0.012	-0.021	-0.170*	0.082*
	(0.027)	(0.061)	(0.045)	(0.023)	(0.092)	(0.046)
Δcurrent rain-March	-0.012	-0.014	0.031**	-0.007	0.036	0.014
	(0.008)	(0.021)	(0.013)	(0.008)	(0.022)	(0.016)
Δcurrent rain-April	-0.033	-0.057	0.041	-0.011	-0.027	0.053
	(0.022)	(0.056)	(0.041)	(0.022)	(0.072)	(0.045)
Δcurrent rain-May	-0.065**	-0.187**	0.107**	-0.031	0.104	-0.023
	(0.027)	(0.073)	(0.043)	(0.026)	(0.095)	(0.052)
Δcurrent rain-June	0.029*	0.092**	-0.047*	0.007	-0.022	-0.094***
	(0.015)	(0.038)	(0.025)	(0.015)	(0.043)	(0.029)
Δcurrent rain-July	-0.014	-0.016	0.036	0.006	0.101	-0.033
	(0.019)	(0.055)	(0.028)	(0.017)	(0.074)	(0.035)
Δcurrent rain-August	-0.021	-0.111	0.010	-0.012	-0.017	-0.108*
	(0.030)	(0.079)	(0.056)	(0.031)	(0.100)	(0.059)
Δcurrent rain-September	0.009	-0.017	-0.028	0.006	0.057	-0.153***
	(0.024)	(0.063)	(0.037)	(0.022)	(0.057)	(0.049)
Δcurrent rain-October	-0.002	0.085	-0.038	0.015	-0.115	0.057
	(0.025)	(0.060)	(0.049)	(0.025)	(0.072)	(0.051)
Δcurrent rain-November	-0.008	-0.086**	0.079**	-0.005	-0.027	0.060*
	(0.019)	(0.040)	(0.035)	(0.018)	(0.057)	(0.031)
Δcurrent rain-December	0.030	-0.114	0.082	-0.003	-0.173	0.048
	(0.038)	(0.100)	(0.055)	(0.035)	(0.108)	(0.066)

Table 1.13 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Δcurrent temperature-Jan	0.142***	0.220	0.027	0.005	-0.261	-0.107
	(0.055)	(0.137)	(0.093)	(0.054)	(0.167)	(0.115)
Δcurrent temperature-Feb	-0.013	-0.408*	0.354**	-0.048	-0.455	-0.034
	(0.093)	(0.246)	(0.156)	(0.088)	(0.309)	(0.186)
Δcurrent temperature-Mar	-0.208**	-0.295	0.369**	-0.134	0.274	0.532**
	(0.096)	(0.216)	(0.169)	(0.090)	(0.209)	(0.208)
Δcurrent temperature-Apr	-0.029	-0.077	0.144**	-0.069*	-0.071	0.115
	(0.044)	(0.092)	(0.073)	(0.040)	(0.102)	(0.089)
Δcurrent temperature-May	0.044	0.044	-0.031	0.008	0.054	-0.164
	(0.067)	(0.147)	(0.106)	(0.065)	(0.219)	(0.124)
Δcurrent temperature-Jun	0.003	-0.006	0.129**	0.004	-0.057	-0.063
	(0.034)	(0.079)	(0.062)	(0.034)	(0.090)	(0.069)
Δcurrent temperature-Jul	-0.050	0.007	0.014	0.001	0.188*	0.013
	(0.032)	(0.092)	(0.050)	(0.029)	(0.100)	(0.061)
Δcurrent temperature-Aug	0.004	-0.013	0.089	-0.018	-0.129	0.107
	(0.040)	(0.113)	(0.062)	(0.039)	(0.108)	(0.082)
Δcurrent temperature-Sep	0.073	-0.056	0.146	0.042	-0.237	-0.303*
	(0.081)	(0.201)	(0.134)	(0.083)	(0.253)	(0.183)
Δcurrent temperature-Oct	0.141	0.308	-0.333*	0.069	0.081	-0.453**
	(0.103)	(0.192)	(0.179)	(0.095)	(0.199)	(0.212)
Δcurrent temperature-Nov	-0.073	-0.224*	0.185*	-0.059	0.107	0.144
	(0.063)	(0.129)	(0.103)	(0.058)	(0.154)	(0.114)
Δcurrent temperature-Dec	0.073	-0.108	-0.086	0.019	0.165	-0.196
	(0.054)	(0.126)	(0.106)	(0.048)	(0.138)	(0.123)
Δlog price of most eaten grain in	0.022	0.076**	-0.015	0.008	-0.078	-0.040
community (rice, flour, or corn)	(0.015)	(0.037)	(0.026)	(0.015)	(0.053)	(0.028)
share males 0-2 yrs old	0.026	0.255	-0.209	-0.152**	0.112	-0.196
	(0.085)	(0.314)	(0.206)	(0.075)	(0.264)	(0.193)
share males 3-5 yrs old	-0.093	0.238	-0.152	-0.228***	-0.400**	-0.121
	(0.067)	(0.164)	(0.148)	(0.069)	(0.187)	(0.190)
share males 6-8 yrs old	-0.056	0.062	-0.228	-0.148**	-0.028	-0.283*
	(0.065)	(0.181)	(0.142)	(0.060)	(0.217)	(0.167)
share males 9-11 yrs old	-0.018	0.053	-0.115	-0.226***	-0.175	-0.047
	(0.065)	(0.158)	(0.143)	(0.061)	(0.180)	(0.160)
share males 12-14 yrs old	-0.021	0.067	-0.131	-0.158***	0.019	-0.033
	(0.059)	(0.237)	(0.129)	(0.059)	(0.277)	(0.168)
share males 15-17 yrs old	-0.004	0.181	-0.023	-0.127**	0.067	0.012
	(0.052)	(0.176)	(0.121)	(0.054)	(0.250)	(0.172)
share males 18-24 yrs old	-0.002	0.132	0.072	-0.032	-0.306***	-0.039
	(0.046)	(0.107)	(0.124)	(0.046)	(0.116)	(0.165)
share males 51-59 yrs old	-0.027	-0.493	0.068	-0.031	-0.066	-0.011
	(0.068)	(0.404)	(0.139)	(0.050)	(0.155)	(0.124)
share males 60+ yrs old	-0.090	-0.136	-0.021	-0.063	-0.089	0.188
	(0.066)	(0.396)	(0.134)	(0.054)	(0.100)	(0.143)
share females 0-2 yrs old	0.029	0.256	0.145	-0.069	-0.054	0.072
	(0.089)	(0.218)	(0.154)	(0.076)	(0.205)	(0.217)

Table 1.13 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(A) Main-Equation Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
share females 3-5 yrs old	-0.018	-0.047	-0.058	-0.181***	0.010	-0.096
	(0.061)	(0.368)	(0.136)	(0.062)	(0.192)	(0.160)
share females 6-8 yrs old	0.009	-0.076	-0.238*	-0.148**	-0.038	0.005
	(0.064)	(0.285)	(0.144)	(0.066)	(0.246)	(0.169)
share females 9-11 yrs old	-0.117*	-0.065	0.042	-0.207***	0.029	0.014
	(0.070)	(0.221)	(0.128)	(0.060)	(0.255)	(0.167)
share females 12-14 yrs old	0.029	0.450	-0.142	-0.172***	0.206	-0.135
	(0.061)	(0.283)	(0.126)	(0.057)	(0.241)	(0.158)
share females 15-17 yrs old	0.056	0.027	-0.074	-0.072	-0.037	-0.026
	(0.062)	(0.188)	(0.126)	(0.053)	(0.220)	(0.148)
share females 18-24 yrs old	-0.032	0.173	-0.158	-0.088*	-0.124	-0.128
	(0.056)	(0.115)	(0.130)	(0.050)	(0.116)	(0.161)
share females 25-50 yrs old	-0.019	0.114	-0.069	-0.091	-0.104	0.068
	(0.067)	(0.111)	(0.181)	(0.066)	(0.140)	(0.224)
share females 51-59 yrs old	0.007	0.146	0.137	-0.083	0.401	0.006
	(0.072)	(0.154)	(0.166)	(0.078)	(0.366)	(0.242)
share females 60+ yrs old	-0.057	0.194	0.125	-0.201***	0.033	0.085
	(0.066)	(0.150)	(0.140)	(0.068)	(0.256)	(0.159)
log household size	-0.005	-0.042	0.013	0.015	-0.001	0.010
(no change in hh size b/w 91 & 93)	(0.014)	(0.113)	(0.034)	(0.014)	(0.074)	(0.031)
Liaoning province (dummy)	-0.152	-0.195	0.126	-0.094	-0.133	0.258
	(0.140)	(0.350)	(0.219)	(0.131)	(0.408)	(0.242)
Henan province (dummy)	0.070	-0.038	0.064	0.019	-0.189	-0.270*
	(0.073)	(0.215)	(0.129)	(0.070)	(0.232)	(0.147)
Shandong province (dummy)	0.041	0.079	-0.208**	0.014	0.041	-0.320***
	(0.058)	(0.173)	(0.100)	(0.057)	(0.183)	(0.106)
Hubei province (dummy)	0.005	-0.214	0.142	-0.012	-0.100	-0.196
	(0.083)	(0.245)	(0.136)	(0.078)	(0.260)	(0.163)
Hunan province (dummy)	0.113	-0.384	0.138	-0.030	-0.197	-0.234
	(0.104)	(0.304)	(0.170)	(0.098)	(0.334)	(0.204)
Guangxi province (dummy)	0.373*	-0.208	0.460	-0.029	-0.785	-0.561
	(0.202)	(0.520)	(0.352)	(0.196)	(0.528)	(0.452)
Guizhou province (dummy)	0.035	-0.596*	0.559**	-0.051	0.311	-0.064
	(0.150)	(0.347)	(0.241)	(0.149)	(0.415)	(0.287)
Constant	0.321	1.048*	-0.633	0.282	0.356	-0.630
	(0.242)	(0.543)	(0.434)	(0.213)	(0.601)	(0.406)
Sample size	2633	450	1728	2929	496	1521
p-value of Hansen J statistic (over-identification test statistic)	0.93	0.22	0.33	0.78	0.30	0.50

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.
2) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.14: 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
F-statistic on the excluded IV's	F(6,2248) = 6.28	F(6, 447) = 3.38	F(6,1410) = 3.82	F(6,2511) = 7.42	F(6, 489) = 2.18	F(6,1219) = 3.01
p-value joint significance	0.0000	0.0029	0.0009	0.0000	0.0437	0.0064
Δprevious rain-February	-0.213***	-0.449***	-0.179***	-0.210***	-0.335***	-0.149**
	(0.049)	(0.113)	(0.065)	(0.048)	(0.114)	(0.075)
Δprevious rain-May	-0.305***	-0.421***	-0.263***	-0.293***	-0.342**	-0.245***
	(0.064)	(0.150)	(0.084)	(0.063)	(0.134)	(0.088)
Δprevious rain-July	0.168***	0.060	0.167**	0.209***	0.160	0.253***
	(0.059)	(0.113)	(0.079)	(0.056)	(0.108)	(0.080)
Δprevious rain-August	0.137*	0.070	0.218*	0.150**	0.095	0.219*
	(0.079)	(0.189)	(0.113)	(0.077)	(0.186)	(0.117)
Δprevious rain-September	0.045	-0.026	0.019	0.037	0.128	0.059
	(0.039)	(0.093)	(0.049)	(0.040)	(0.091)	(0.057)
Δprevious rain-December	0.300***	0.617***	0.267**	0.270***	0.349**	0.248**
	(0.081)	(0.192)	(0.109)	(0.076)	(0.160)	(0.112)
male 0-2 yrs old (dummy)			-0.032			
			(0.042)			
male 3-5 yrs old (dummy)			0.007			
			(0.036)			
male 6-8 yrs old (dummy)			-0.001			
			(0.030)			
male 9-11 yrs old (dummy)			-0.011			
			(0.027)			
male 12-14 yrs old (dummy)			-0.009			
			(0.024)			
male 18-20 yrs old (dummy)	0.032					
	(0.029)					
male 21-23 yrs old (dummy)	0.021					
	(0.031)					
male 24-26 yrs old (dummy)	0.045					
	(0.030)					
male 27-29 yrs old (dummy)	0.039					
	(0.026)					
male 33-35 yrs old (dummy)	0.031					
	(0.024)					
male 36-38 yrs old (dummy)	0.006					
	(0.027)					
male 39-41 yrs old (dummy)	0.046*					
	(0.026)					
male 42-44 yrs old (dummy)	0.046					
	(0.029)					
male 45-47 yrs old (dummy)	0.051					
	(0.032)					
male 48-50 yrs old (dummy)	0.040					
	(0.031)					

Table 1.14 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
male 51-53 yrs old (dummy)	0.041 (0.036)					
male 54-56 yrs old (dummy)	0.035 (0.031)					
male 57-59 yrs old (dummy)	0.072** (0.030)					
male 63-65 yrs old (dummy)		0.030 (0.038)				
male 66-68 yrs old (dummy)		-0.025 (0.044)				
male 69-71 yrs old (dummy)		-0.035 (0.049)				
male 72-74 yrs old (dummy)		-0.013 (0.045)				
male 75-77 yrs old (dummy)		-0.037 (0.063)				
male 78-80 yrs old (dummy)		-0.114* (0.069)				
male 81-83 yrs old (dummy)		-0.088 (0.093)				
male 84+ yrs old (dummy)		0.028 (0.148)				
female 0-2 yrs old (dummy)						0.021 (0.039)
female 3-5 yrs old (dummy)						-0.001 (0.032)
female 6-8 yrs old (dummy)						0.011 (0.030)
female 9-11 yrs old (dummy)						-0.003 (0.028)
female 12-14 yrs old (dummy)						0.023 (0.024)
female 18-20 yrs old (dummy)				0.037 (0.027)		
female 21-23 yrs old (dummy)				0.058* (0.031)		
female 24-26 yrs old (dummy)				0.021 (0.027)		
female 27-29 yrs old (dummy)				0.034 (0.024)		
female 33-35 yrs old (dummy)				0.030 (0.024)		
female 36-38 yrs old (dummy)				0.056** (0.025)		

Table 1.14 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
female 39-41 yrs old (dummy)				0.084***		
				(0.026)		
female 42-44 yrs old (dummy)				0.064**		
				(0.028)		
female 45-47 yrs old (dummy)				0.076**		
				(0.031)		
female 48-50 yrs old (dummy)				0.039		
				(0.030)		
female 51-53 yrs old (dummy)				0.035		
				(0.031)		
female 54-56 yrs old (dummy)				0.048		
				(0.031)		
female 57-59 yrs old (dummy)				0.014		
				(0.031)		
female 63-65 yrs old (dummy)					0.036	
					(0.035)	
female 66-68 yrs old (dummy)					0.016	
					(0.041)	
female 69-71 yrs old (dummy)					0.005	
					(0.042)	
female 72-74 yrs old (dummy)					0.041	
					(0.050)	
female 75-77 yrs old (dummy)					0.090*	
					(0.048)	
female 78-80 yrs old (dummy)					0.081	
					(0.049)	
female 81-83 yrs old (dummy)					0.141**	
					(0.064)	
female 84+ yrs old (dummy)					0.064	
					(0.059)	
ever farm (dummy)	0.021	-0.064	0.003	0.022	0.044	0.045*
	(0.018)	(0.044)	(0.022)	(0.017)	(0.039)	(0.026)
village resident (dummy)	-0.070***	0.019	-0.019	-0.042**	-0.032	-0.055**
	(0.020)	(0.049)	(0.025)	(0.019)	(0.043)	(0.027)
Δcurrent rain-January	-0.336***	-0.640***	-0.280***	-0.315***	-0.402**	-0.227**
	(0.071)	(0.186)	(0.090)	(0.064)	(0.158)	(0.103)
Δcurrent rain-February	-0.340***	-0.436***	-0.316***	-0.314***	-0.488***	-0.267***
	(0.060)	(0.129)	(0.085)	(0.059)	(0.123)	(0.089)
Δcurrent rain-March	-0.010	-0.092**	-0.021	-0.012	0.007	-0.004
	(0.021)	(0.044)	(0.029)	(0.020)	(0.040)	(0.029)
Δcurrent rain-April	-0.022	0.068	-0.050	-0.033	-0.168*	-0.029
	(0.040)	(0.088)	(0.054)	(0.039)	(0.090)	(0.056)
Δcurrent rain-May	0.301***	0.504***	0.300**	0.320***	0.691***	0.318**
	(0.092)	(0.193)	(0.119)	(0.092)	(0.192)	(0.133)

Table 1.14 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Δcurrent rain-June	0.001	-0.124	0.020	0.011	-0.081	0.010
	(0.033)	(0.076)	(0.044)	(0.031)	(0.067)	(0.048)
Δcurrent rain-July	0.094**	0.072	0.104**	0.100**	0.301***	0.098*
	(0.040)	(0.089)	(0.052)	(0.039)	(0.083)	(0.059)
Δcurrent rain-August	-0.152***	-0.350**	-0.173**	-0.138***	-0.140	-0.102
	(0.053)	(0.140)	(0.069)	(0.051)	(0.137)	(0.077)
Δcurrent rain-September	0.077*	0.066	0.062	0.085**	0.113	0.151***
	(0.042)	(0.091)	(0.052)	(0.041)	(0.079)	(0.058)
Δcurrent rain-October	0.052	0.128	0.126	0.093	0.031	0.159*
	(0.062)	(0.153)	(0.081)	(0.061)	(0.161)	(0.085)
Δcurrent rain-November	-0.197***	-0.344***	-0.168**	-0.183***	-0.397***	-0.104
	(0.053)	(0.118)	(0.069)	(0.049)	(0.112)	(0.081)
Δcurrent rain-December	-0.179**	-0.353**	-0.201*	-0.221***	-0.253	-0.079
	(0.081)	(0.175)	(0.105)	(0.077)	(0.163)	(0.111)
Δcurrent temperature-Jan	-0.542***	-0.709*	-0.422*	-0.600***	-1.076***	-0.437
	(0.190)	(0.361)	(0.251)	(0.184)	(0.373)	(0.266)
Δcurrent temperature-Feb	0.967***	1.789**	1.238***	1.011***	1.160	1.369***
	(0.320)	(0.778)	(0.442)	(0.321)	(0.711)	(0.447)
Δcurrent temperature-Mar	-0.601***	-0.866**	-0.467**	-0.548***	-0.377	-0.721***
	(0.180)	(0.367)	(0.217)	(0.172)	(0.343)	(0.239)
Δcurrent temperature-Apr	-0.424***	-0.627***	-0.325***	-0.382***	-0.575***	-0.388***
	(0.083)	(0.198)	(0.102)	(0.081)	(0.190)	(0.112)
Δcurrent temperature-May	0.450***	0.316	0.449**	0.504***	0.938***	0.441**
	(0.140)	(0.292)	(0.177)	(0.133)	(0.291)	(0.199)
Δcurrent temperature-Jun	0.279***	0.375*	0.277**	0.275***	0.369**	0.379***
	(0.086)	(0.197)	(0.111)	(0.082)	(0.180)	(0.124)
Δcurrent temperature-Jul	0.307***	0.398**	0.276**	0.282***	0.501***	0.202*
	(0.081)	(0.174)	(0.109)	(0.076)	(0.163)	(0.115)
Δcurrent temperature-Aug	-0.453***	-0.592***	-0.398***	-0.515***	-0.539**	-0.559***
	(0.103)	(0.221)	(0.128)	(0.099)	(0.211)	(0.140)
Δcurrent temperature-Sep	-0.104	-0.314	-0.016	-0.129	-0.534	0.285
	(0.237)	(0.511)	(0.340)	(0.232)	(0.475)	(0.363)
Δcurrent temperature-Oct	0.824***	0.991***	0.746***	0.746***	0.790**	0.773***
	(0.163)	(0.348)	(0.192)	(0.150)	(0.326)	(0.219)
Δcurrent temperature-Nov	-0.054	-0.086	0.073	0.041	0.170	0.045
	(0.123)	(0.269)	(0.140)	(0.115)	(0.250)	(0.157)
Δcurrent temperature-Dec	0.044	0.012	0.005	0.036	0.136	0.144
	(0.136)	(0.315)	(0.190)	(0.132)	(0.336)	(0.194)
Δlog price of most eaten grain in community (rice, flour, or corn)	0.028	0.026	0.021	0.039	-0.074	0.026
	(0.031)	(0.061)	(0.041)	(0.030)	(0.065)	(0.043)
share males 0-2 yrs old	-0.021	-0.294	0.016	-0.009	-0.079	0.310
	(0.128)	(0.323)	(0.234)	(0.121)	(0.251)	(0.225)
share males 3-5 yrs old	0.165	0.066	-0.062	0.184*	0.098	0.498**
	(0.101)	(0.225)	(0.192)	(0.098)	(0.201)	(0.197)
share males 6-8 yrs old	0.200*	0.357	-0.055	0.112	0.248	0.366*
	(0.103)	(0.292)	(0.170)	(0.100)	(0.305)	(0.204)

Table 1.14 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
share males 9-11 yrs old	0.087	0.100	-0.093	0.028	0.046	0.363*
	(0.102)	(0.275)	(0.168)	(0.101)	(0.213)	(0.199)
share males 12-14 yrs old	0.076	0.592**	-0.049	0.041	0.545**	0.230
	(0.105)	(0.293)	(0.164)	(0.104)	(0.218)	(0.201)
share males 15-17 yrs old	-0.025	-0.057	-0.073	-0.060	-0.344	0.304
	(0.105)	(0.254)	(0.158)	(0.099)	(0.254)	(0.201)
share males 18-24 yrs old	0.044	0.212	-0.068	-0.089	-0.011	0.280
	(0.085)	(0.152)	(0.157)	(0.084)	(0.133)	(0.184)
share males 51-59 yrs old	-0.111	-0.326	-0.249	-0.011	0.108	-0.017
	(0.100)	(0.473)	(0.200)	(0.092)	(0.202)	(0.264)
share males 60+ yrs old	-0.112	0.059	-0.124	-0.057	-0.171	0.099
	(0.114)	(0.429)	(0.163)	(0.096)	(0.123)	(0.202)
share females 0-2 yrs old	0.020	0.084	0.125	0.006	0.073	0.318
	(0.124)	(0.288)	(0.204)	(0.124)	(0.302)	(0.226)
share females 3-5 yrs old	-0.071	0.426	-0.100	-0.083	0.197	0.157
	(0.104)	(0.365)	(0.163)	(0.102)	(0.319)	(0.198)
share females 6-8 yrs old	0.033	-0.291	-0.030	0.058	0.048	0.229
	(0.104)	(0.306)	(0.155)	(0.100)	(0.248)	(0.190)
share females 9-11 yrs old	0.024	0.134	-0.096	-0.026	0.205	0.233
	(0.105)	(0.363)	(0.156)	(0.104)	(0.265)	(0.200)
share females 12-14 yrs old	0.121	0.480	-0.020	0.008	0.353	0.179
	(0.110)	(0.324)	(0.157)	(0.104)	(0.259)	(0.208)
share females 15-17 yrs old	-0.105	-0.333	-0.157	-0.178*	0.081	0.118
	(0.114)	(0.240)	(0.171)	(0.098)	(0.228)	(0.187)
share females 18-24 yrs old	-0.089	-0.068	-0.191	-0.083	-0.018	0.170
	(0.097)	(0.173)	(0.156)	(0.092)	(0.176)	(0.213)
share females 25-50 yrs old	-0.017	0.021	-0.135	0.046	-0.114	0.398
	(0.126)	(0.206)	(0.209)	(0.121)	(0.200)	(0.264)
share females 51-59 yrs old	0.025	-0.383*	-0.127	0.096	-0.536	0.524*
	(0.132)	(0.201)	(0.207)	(0.132)	(0.360)	(0.316)
share females 60+ yrs old	0.089	-0.300	-0.021	-0.018	0.248	0.206
	(0.115)	(0.187)	(0.169)	(0.107)	(0.286)	(0.211)
log household size	-0.002	-0.040	0.019	0.007	0.018	-0.012
(no change in hh size b/w 91 & 93)	(0.023)	(0.110)	(0.034)	(0.022)	(0.079)	(0.044)
Liaoning province (dummy)	-1.549***	-2.087**	-1.078**	-1.364***	-1.864**	-1.183**
	(0.359)	(1.037)	(0.473)	(0.351)	(0.894)	(0.560)
Henan province (dummy)	-0.138	-0.294	-0.011	-0.107	-0.265	0.024
	(0.174)	(0.450)	(0.229)	(0.165)	(0.365)	(0.249)
Shandong province (dummy)	0.165	-0.090	0.197	0.164	0.155	0.243
	(0.137)	(0.381)	(0.181)	(0.132)	(0.303)	(0.192)
Hubei province (dummy)	-0.033	-0.330	0.154	0.029	-0.026	0.126
	(0.196)	(0.512)	(0.265)	(0.188)	(0.432)	(0.283)
Hunan province (dummy)	-0.278	-0.808	-0.046	-0.168	-0.350	0.042
	(0.218)	(0.613)	(0.284)	(0.208)	(0.497)	(0.303)

Table 1.14 (continued): 2SLS Regressions of Individual Caloric Intakes (Full Results)
(B) First-Stage Results

demographic group	(1) M prime-age	(2) M elderly	(3) M children	(4) F prime-age	(5) F elderly	(6) F children
Guangxi province (dummy)	0.053	-0.934	0.693	0.227	-0.945	1.191*
	(0.462)	(1.036)	(0.659)	(0.453)	(0.939)	(0.689)
Guizhou province (dummy)	0.212	0.050	0.467	0.323	0.975	0.453
	(0.297)	(0.704)	(0.361)	(0.280)	(0.640)	(0.400)
Constant	0.081	-0.322	-0.184	-0.354	-0.073	-0.549
	(0.453)	(1.022)	(0.639)	(0.431)	(1.000)	(0.680)
Sample size	2633	450	1728	2929	496	1521

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.

2) * significant at 10%; ** significant at 5%; *** significant at 1%

References

Alderman, Harold, and Paul Gertler. (1997). "Family Resources and Gender Differences in Human Capital Investments: The Demand for Children's Medical Care in Pakistan," in *Intrahousehold Resource Allocation in Developing Countries: Models, Methods, and Policy*, Haddad, Hoddinott, and Alderman eds., Baltimore and London: The Johns Hopkins University Press.

Behrman, Jere R. (1988). "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?" *Oxford Economic Papers*, 40 (1): 32-54.

Behrman, Jere R., and Anil Deolalikar. (1990). "The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects and Permanent Income," *The Journal of Human Resources*, 25: 655-696.

Carolina Population Center, The University of North Carolina at Chapel Hill. Further information of the CHNS is available at www.cpc.unc.edu/projects/china/china_home.html (accessed on July 31, 2004).

Chinese Nutrition Society. (2001). *Chinese DRIs*, in Chinese, ISBN 7-5019-3147-X.

Coale, Ansley. J. (1991). "Excess Female Mortality and the Balance of the Sexes in the Population: An Estimate of the Number of Missing Females," *Population and Development Review*, 17 (3): 517-523.

de Tray, Dennis. (1988). "Government Policy, Household Behavior, and the Distribution of Schooling: A Case Study of Malaysia," in *Research in Population Economics 6*, T. Paul Schultz ed., Greenwich, Conn. and London: JAI Press.

Dercon, Stefan, and Pramila Krishnan. (2000). "In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia," *Journal of Political Economy*, 108(4): 688-727.

Fafchamps, Marcel. (1993). "Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers," *Econometrica*, 61 (5): 1173-97.

Gertler, Paul, and Paul Glewwe. (1992). "The Willingness to Pay for Education for Daughters in Contrast to Sons: Evidence from Rural Peru," *The World Bank Economic Review*, 6(1): 171-88.

Giles, John. (2003). "Beg, Borrow or Dis-Save? Mechanisms Used to Cope with Unexpected Shocks to Income in Rural China," mimeo, Michigan State University.

Jalan, Jyotsna and Martin Ravallion. (1999). "Are the Poor Less Well Insured? Evidence on Vulnerability to Income Risk in Rural China," *Journal of Development Economics*, 58 (1): 61-81.

Jamison, Dean T. (1986). “Child Malnutrition and School Performance in China,” *Journal of Development Economics*, 20: 299-309.

Johansson, Sen, and Ola Nygren. (1991). “The Missing Girls of China: A New Demographic Account,” *Population and Development Review*, 17 (1): 35-51.

Junhong, Chu. (2001). “Prenatal Sex Determination and Sex-Selective Abortion in Rural Central China,” *Population and Development Review*, 27 (2): 259-281.

King, Elizabeth M., and Lee A. Lillard. (1987). “Education Policy and Schooling Attainment in Malaysia and the Philippines,” *Economics of Education Review*, 6(2): 167-81.

Kochar, Anjini. (1999a). “Evaluating Familial Support for the Elderly: the Intra-household Allocation of Medical Expenditures in Rural Pakistan,” *Economic Development and Cultural Change*, 47 (3): 620-56.

Kochar, Anjini. (1999b). “Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India,” *Review of Economics and Statistics*, (81) (1): 50-61.

Kwong, P. and G. Cai. (1992). "Ageing in China: Trends, Problems and Strategies," in *Ageing in East and South-east Asia*, David R. Phillips ed., London: E. Arnold.

Lavy, Victor. (1996). "School Supply Constraints and Children's Educational Outcomes in Rural Ghana," *Journal of Development Economics*, 51(2): 291-314.

Lee, Yean-Ju, and Zhenyu Xiao. (1998). "Children's Support for Elderly Parents in Urban and Rural China: Results from a National Survey," *Journal of Cross-Cultural Gerontology*, 13: 39-62.

Lee, James Z., and Wang Feng. (1999). *One Quarter of Humanity: Malthusian Mythology and Chinese Realities,1700-2000*, Cambridge, Massachusetts: Harvard University Press.

Miguel, Edward. (2003). “Poverty and Witch Killing” mimeo, UC Berkeley and NBER.

Morgan, Stephen L. (2000). "Richer and Taller: Stature and Living Standards in China, 1979-1995," *China Journal*, 44: 1-39.

Park, Albert, and Pungpond Rukumnuaykit. (2004). "Eat Drink Man Woman: Testing for Gender Bias in China Using Individual Nutrient Intake Data," mimeo, University of Michigan.

Pei, X. and V.K. Pillai. (1999). "Old Age in China: The Role of the State and the Family," *International Journal of Aging and Human Development*, 49: 197-212.

Pelletier, David L., and Edward A. Frongillo. (2003). "Changes in Child Survival Are Strongly Associated with Changes in Malnutrition in Developing Countries," *Journal of Nutrition*, 133: 107-119.

Pitt, Mark M., Mark R. Rosenzweig, and Nazmul Hassan. (1990). "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries," *The American Economic Review*, 80 (5): 1139-1156.

Rose, Elaina. (1999). "Consumption Smoothing and Excess Female Mortality in Rural India," *Review of Economics and Statistics*, 81(1): 41-49.

Rose, Elaina. (2001). "Ex Ante and Ex Post Labor Supply Response to Risk in a Low-Income Area," *Journal of Development Economics*, 64: 371-388.

Schroeder, Dirk E. and Reynaldo Martorell. (1997). "Enhancing Child Survival by Preventing Malnutrition," *American Journal of Clinical Nutrition*, 65: 1080-81.

Schultz, T. Paul. (1987). "School Expenditures and Enrollments, 1960-1980: The Effects of Income, Prices and Population Growth," in *Population Growth and Economic Development*, D. Gale Johnson and R. Lee, eds., Madison: University of Wisconsin Press.

Sipahimalani, Vandana. (1999). "Education in the Rural Indian Household: The Impact of Household and School Characteristics on Gender Differences," Working Paper 68, National Council of Applied Economic Research, New Delhi.

Skoufias, Emmanuel. (1993). "Seasonal Labor Utilization in Agriculture: Theory and Evidence from Agrarian Households in India," *American Journal of Agricultural Economics*, 75 (1): 20-32.

Tolley, G. S., and R. W. Gieseman. (1963). "Consumer Demand Explained by Measurable Utility Changes," *Econometrica*, 31 (3): 499-513.

World Bank. (2001). *Engendering Development Through Gender Equality in Rights, Resources, and Voice*, World Bank Policy Research Report, World Bank and Oxford University Press.

WHO Global Database on Child Growth and Malnutrition, available online www.who.int/nutgrowthdb/ (accessed on July 31, 2004).

Wiseman Paul. (2002). "China Thrown Off Balance as Boys Outnumber Girls," *USA Today*, June 19, 2002.

Yu, Mei-Yu, and Rosemary Sarri. (1997). "Women's Health Status and Gender Inequality in China," *Social Science & Medicine*, 45 (12): 1885-1898.

Zeng Yi, Tu Ping, Gu Baochang, Xu Yi, Li Bohua, Li Yongpiing. (1993). "Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China," *Population and Development Review*, 19 (2): 283-302.

Chapter 2

Measuring the Impact of Health Insurance on Physician Visits by the Elderly: A Natural Experiment in Taiwan

Abstract

This chapter measures the impact of health insurance on physician visits by the elderly in Taiwan, using panel data that span the introduction of national health insurance (NHI) in 1995. Using four waves of data, (two waves before and two waves after the introduction of NHI), I calculate difference-in-difference estimates of the effect of insurance on whether the elderly make any visit to doctors and the number of visits conditional on visiting a doctor, using linear, discrete-choice, and count-data models. I find that four years after the introduction of NHI, health insurance coverage had no effect on the probability of having at least one physician visit, although it increased conditional physician visits by 29%. I also find evidence that factors such as pent-up demand and congestion of medical facilities may have affected the magnitudes and timing of insurance effects. Finally, using a conventional instrumental variable approach, I find that the impact of health insurance on conditional physician visits is much larger than that estimated using the natural experiment as the source of identification. This may signify the importance of the natural experiment to convincingly estimate the impact of health insurance on health care utilization.

1. Introduction

This chapter measures the impact of health insurance on physician visits by the elderly in Taiwan, using panel data that span the introduction of national health insurance (NHI) in 1995.[1] Because four waves of data are available (two waves before the

[1] For this study, I focus on outpatient visits to western-medicine physicians. Inpatient visits are not included. Visits to pharmacies and Chinese-medicine physicians are not considered either in this chapter.

introduction of NHI in 1989 and 1993 and two waves after in 1996 and 1999), multiple difference-in-difference estimates can be calculated using different combinations of yearly waves.

The availability of data four years after implementation makes it possible to separately estimate short-term and medium-term effects of the program. This is important because short-term effects can be influenced by such factors as pent-up demand for health care and congestion of medical facilities. We find evidence that physician visits one year after the introduction of NHI were likely influenced by the pent-up demand for health care and congestion of medical facilities. Table 2.1 reports medical utilization statistics for Taiwan from 1992 to 2001. Figures 2.1 through 2.3 plot the numbers of outpatient, inpatient, and emergency visits, respectively. While medical utilization generally increased over the period, the changes from 1994 to 1996 are irregular. Outpatient visits decreased from 1994 to 1995 and increased by 8.5% from 1995 to 1996. Inpatient visits increased slightly from 1994 to 1995 (0.90%) and much faster from 1995 to 1996 (5.53%). The number of emergency visits increased by about seven percent both from 1994 to 1995 and from 1995 to 1996. The national statistics suggest that both outpatient and inpatient visits went through a period of adjustment immediately after NHI.

Examining effects four years after NHI was introduced, I find that health insurance coverage had no significant effect on whether patients visited a doctor in the past month, but it increased the number of physician visits in the past month by 29% conditional on seeing a doctor. This chapter also calculates estimates of the impact of health insurance with an instrumental variable (IV) using pre-NHI waves of the data to test the accuracy of methodologies used when a natural experiment is absent.

Taiwan has a history of having virtually no private health insurance. Prior to NHI, public health-insurance programs were established for only a subset of the population. For example, salaried workers,[2] government employees (including retired government employees and family members), farmers (including family members who farm), and

[2] Employers with five or more employees were required by law to insure all workers between the ages of 15 and 60 years, but coverage is not mandatory for dependents. Peabody et al (1995), Chiang (1997), Tsai et al (1998), Chou et al (2001), Chou et al (2002), and Cheng (2003) contain descriptions of social health-insurance programs in Taiwan before 1995.

military-service persons (including veterans and family members) pooled health insurance funds respectively, while the self-employed did not have any public health-insurance program for which they were eligible. Labor Insurance (LI), Government Employees Insurance (GEI), Farmers Insurance (FI), and Low-Income Household Insurance (LIHI) started in 1950, 1958, 1985, and 1990. Spouses, parents, and children of government employees gained coverage under Government Employees Dependents Insurance (GEDI) starting in 1982, 1989, and 1992. Retired government employees and their dependents became eligible for coverage in 1985. The scope of coverage was quite comprehensive and similar across the different public health-insurance programs. The Taiwanese government subsidized different social health-insurance programs to different degrees. Prior to the introduction of NHI, approximately 60 percent of the Taiwanese population were covered by some types of social health insurance. Children, the elderly, and women not employed outside of the home comprised most of the uninsured individuals.

In July 1994, the NHI Law passed the Congress. In March 1995, NHI was brought into effect and replaced all existing social health-insurance programs, and extended coverage to all of Taiwan's citizens. Since NHI is mandatory by law, virtually all Taiwanese citizens enrolled in NHI when it was instituted in March 1995. NHI requires Taiwanese residents to pay insurance premiums whose amount depends upon each individual's income, occupation, and number of dependents. The Taiwanese government subsidizes NHI. For example, the government pays 60% of the premiums for government employees and their dependents, 70% of the premiums for farmers and their dependents, and the entire premiums for low-income households and military personnel and their dependents. The government contributes 10% of the premiums for salaried workers and their dependents; the rest is shared by the insured (30%) and their employers (60%). Under NHI, the insured have the same comprehensive coverage no matter how much they pay for their premiums. The range of coverage is similar to those of the pre-NHI social programs, with some additional coverage for severe illnesses and home health care. NHI does not cover individuals who fail to pay the premiums. As of January 2000, the enrollment rate was 96.13%.[3]

[3] http://www.nhi.gov.tw/00english/e_index.htm (accessed on September 6, 2004)

Measuring the impact of health insurance on physician visits is an important economic question. First, evaluation of universal health insurance in Taiwan requires quantifying the benefits of the policy, and one critical goal of the program was to increase access to physicians by those who were previously uninsured. Further, estimating the impact of health insurance can help policy makers to better understand the sources of inequality in health care utilization. A weak impact of health insurance would suggest that out-of-pocket expenses are not a crucial determinant of health care utilization, implying either that other factors such as time and transportation costs to physicians play a more significant role or that inequality in health care utilization is a matter of taste, preference, or lifestyle. Second, estimates of the increase in demand resulting from NHI may help other countries considering national health insurance to plan appropriately. Third, health is a main concern of the elderly, so promoting greater health care access for the elderly has significant welfare consequences, and could help reduce social inequality, especially in a country like Taiwan that has a rapidly aging population.

Two distinct concepts of the impact of health insurance, one horizontal and one vertical, are measured in this study. The horizontal measure focuses on how broadly health care becomes available due to NHI. The vertical measure examines whether health insurance changes the intensity of physician visits.

To measure the impact of health insurance on physician visits, we employ difference-in-differences estimators to evaluate the impact of NHI on elderly physician visits. Adopting the language used in the program evaluation literature, individuals who were covered both before and after the introduction of NHI serve as the control group, while individuals who were not covered before NHI but were covered after NHI serve as the treatment group. Although there were minor differences in the level of insurance coverage among different types of social health insurance prior to NHI, I treat the social health-insurance programs before NHI as homogenous in the subsequent analysis. Further, I assume that NHI is identical to the public health insurance programs prior to NHI in terms of practical scope of coverage.[4] To test the robustness of estimates that do not benefit from a natural experiment, I also employ another identification strategy that

---

[4] See literature listed in footnote 2.

exploits the variation in health-insurance coverage across individuals before NHI using an instrumental variables approach.

2. Previous Research

Cheng et al (1997) looked at the impact of NHI on health care utilization in Taiwan, using a simple difference-in-differences approach using panel data collected in October through December in 1994 and in December 1995. They concluded that universal health insurance removed some barriers to health care for the newly insured. Specifically, they found that after the implementation of NHI, the newly insured consumed more than twice the amounts of outpatient visits in the two weeks prior to the survey interview date (0.21 versus 0.48) and hospital admissions in the immediate past year (0.04 versus 0.11) than before NHI was introduced, bringing them on a par with the previously insured in terms of health care utilization. Given the timing of the follow-up survey, however, Cheng et al (1997) can only look at short-term effects of NHI.

The Rand randomized experiment by Manning et al (1987) is the best known study of the effect of health insurance on health care utilization. The authors found that the co-payment rate (free, 25%, 50%, 95%) is highly correlated with both the likelihood of any outpatient visit and outpatient expenses conditional on any visit. Since the study randomly assigned different insurance plans to randomly-chosen subjects, the orthogonality between the insurance plans and unobserved determinants of the demand for health care is guaranteed. My study provides two pieces of information that the Rand randomized study does not address. First, the Rand randomized experiment does not include the elderly although elderly utilization of health care is important qualitatively and quantitatively. Second, the context in which national health insurance is introduced in a less-industrialized country may be quite different from the US. Table 2.2 reports health-related indicators for selected countries. Taiwan was an upper middle-income country in 1990.

McWilliams et al (2003), analyzing panel data constructed from the Health and Retirement Study, find that the effect of Medicare coverage on the use of covered clinical examinations (cholesterol testing, mammography, and prostate examination) is substantially larger for the previously uninsured than for the previously insured. They also find that the effect of Medicare coverage on the use of cholesterol testing is significantly larger for the previously uninsured with either hypertension or diabetes than for the previously uninsured without those conditions, implying that adults in greater need of cardiovascular risk reduction benefit more from Medicare coverage than adults in less need. Buchmueller et al (2004) find that hospital closures (which increased distances to the nearest hospital) in Los Angeles County decreased the probability that the elderly received flu shots and decreased the reported ease of access to health care services among lower-income residents. Even in an urban setting, the physical distance to a health-care provider could affect the use of some medical services especially among vulnerable populations such as seniors and lower-income residents. Although universal health insurance in Taiwan would be helpful to reduce the financial barriers to access to health care, especially for the poor, there is a possibility that other barriers such as time and transportation costs could still prevent poor or unhealthy individuals from utilizing health care services.

3. Data

This study uses data from a nationally representative panel dataset originally including 4049 elderly individuals in 1989. The Surveys of Health and Living Status of the Elderly in Taiwan (SHLSE) were conducted in 1989, 1993, 1996, and 1999. The ages of sample individuals in 1989 ranged from 60 to 96.

One attractive feature of the dataset is the exceptionally high response rate, which is especially important for elderly surveys because non-respondents tend to be less healthy than average. The 4049 interviews completed in 1989 constitute a response rate of 91.8%, which was made possible by seriously tracing selected respondents who no longer resided at the sampled registered addresses and by utilizing proxy interviews with household members most knowledgeable about the health and current situations of the respondents with severe infirmity. Response rates are quite even across genders (91.1% for males and 92.7% for females) and age groups (90.8% for those aged 60 to 64, 93.0% for those aged 65 to 69, 93.1% for those aged 70 to 74, 92.5% for those aged 75 to 79, and 87.5% for those aged 80+). Table 2.3 compares the age and gender distributions between the population and the completed sample for each sample year.[5] Panel attrition reduced the total number of individuals to 3155 (77.9%) in 1993, 2669 (65.9%) in 1996, and 2310 (57.1%) in 1999, respectively. Comparing the population distribution with the sample distribution reveals that the subpopulation aged 75 or older and the female population are slightly underrepresented by the sample. Attrition bias due to old age appears minimal. For a detailed description of the survey, see Taiwan Provincial Institute of Family Planning, and Population Studies Center, University of Michigan (1989) and (1997).

4. Descriptive Evidence

To get a better sense of the types of individuals affected by NHI, we first present some descriptive statistics on the insured and uninsured in 1993 before NHI was implemented. Table 2.4 compares the characteristics of individuals who were covered both in 1993 and 1996 (the previously insured) with those who were not covered in 1993 but were covered in 1996 (the previously uninsured). Table 2.4 reports the means and standard deviations of some observable characteristics by insurance status in 1993 as well as the p-values for the mean difference tests on those characteristics. The numbers of the previously insured and uninsured are 1915 and 545, respectively. The previously insured have higher proportions or means than the previously uninsured with respect to the following observable characteristics: the proportion of males (p-value for the mean difference test, 0.000), years of schooling (0.000), marital status (0.000), working status (0.001), the proportion of individuals who have received pension payments (0.000), the proportions of individuals who have real estate (0.000) and savings/stocks (0.001), the proportion of individuals who live alone (0.000), and the proportion of rural residents (0.000). The previously uninsured are on average less healthy than the previously insured in terms of a typical ADL score (0.012). The previously insured and uninsured are likely to be significantly different in unobservable characteristics as well. The inclusion of individual fixed effects facilitated by panel data thus is useful for consistently estimating the impact of health insurance.

[5] For the completed sample, all individuals interviewed (regardless of whether or not she/he answered a particular question) are reported in Table 3.

In Table 2.5, a row labeled "Physician Visits" reports the numbers of mean physician visits in the past month in each survey year separately for the previously insured and uninsured. Standard deviations are in parentheses. Table 2.5 uses only individuals who are present in all sample years. Tracing the same individuals over time, any changes in the number of physician visits for each group should reflect only time-varying determinants of physician visits. "The insured" refers to those with health insurance in all sample years, while "the uninsured" refers to those without health insurance in 1989 and 1993 but with health insurance in 1996 and 1999. The sample sizes for the insured and the uninsured are 1018 and 238.

Figure 2.4 graphs the trends of physician visits separately for the insured and the uninsured, and Figure 2.5 plots the difference in physician visits between the two groups. Physician visits increased for both groups through the sample period. Before NHI, the insured visited physicians more frequently than the uninsured, while after NHI, the uninsured increased visits to the level of the insured, a result consistent with Cheng et al (1997). The mean number of physician visits is larger for the uninsured than for the insured in 1996, but the reverse is true in 1999.

We next define "any visit" to be an indicator variable for whether or not the respondent visited a doctor at least once in the past month preceding the data of the interview, and "conditional visits" to be the number of visits minus one for those who made at least one visit to a doctor.[6] The rows labeled "Any Visit" and "Conditional Visits" of Table 2.5 report the trends for any visit and conditional visits, respectively. The corresponding graphs are presented in Figures 2.6 through 2.9. Any visit equals one if the individual visited a physician once or more in the past month preceding the date of the interview, and zero otherwise. Any visit increased for both groups throughout the sample period. The uninsured were less likely to have visited a physician in the past month in comparison with the insured prior to NHI. After the introduction of NHI, the uninsured were slightly more likely to visit a physician in 1996, but the difference in any visit between the two groups in 1999 returned back to the level of 1989. Conditional visits steadily increased for the uninsured and more or less remained the same for the insured.

[6] One is subtracted for conditional visits because this avoids truncation problems in the econometric analysis of Poisson count data models.

In particular, the insured reduced conditional visits from 1993 to 1996. Even without changes in health insurance status, the uninsured increased conditional visits by 83% from 1989 to 1993 and surpassed the number of conditional visits of the insured.

Physician visits and conditional visits do not count outliers whose physician visits in adjacent sample years differ by more than 20 in absolute value. For example, a report of 22 physician visits in 1989 and one physician visit in 1993 is excluded in calculating the means in Table 2.5. This is done because the sample size of the uninsured for conditional visits is only 27, and I want to avoid the group means being affected by a single individual with a large change in conditional visits. The bottom panel of Table 2.5 shows that including such individuals, however, does not significantly change the results described above, except for the mean conditional visits by the uninsured in 1993 (3.000 with the inclusion of such individuals instead of 2.037).

Because two waves of data on health care utilization are available prior to NHI, we can compare the trend of health care utilization prior to NHI separately for the previously insured and the previously uninsured. This may provide some indication of the counter factual trend of health care utilization after 1995 in the fictitious world where there were no NHI. Table 2.6 shows the mean changes in physician visits, any visit, and conditional visits between 1989 and 1993 separately for the previously insured and the previously uninsured. The previously insured are those with health insurance both in 1989 and 1993, and the previously uninsured are those without health insurance both in 1989 and 1993. Table 2.6 also reports the p-values of the mean difference tests of changes in health care utilization by the two groups.

For physician visits, the mean change is more positive for the previously insured than for the previously uninsured, but the difference is not statistically significant at the conventional significance levels. Changes in any visit are on average more positive for the previously insured than for the previously uninsured, and the difference is statistically significant at the one percent significance level. Changes in conditional visits by the previously insured are on average slightly negative, while changes in conditional visits by the previously uninsured are on average positive. Due to the large standard errors, however, the mean difference is not statistically significant at the ten percent level. Including the outliers, the mean change in conditional visits by the previously uninsured is larger (0.940 instead of 0.602), and the test shows a statistically significant mean difference between the two groups at the ten percent significance level.

If this is indicative of the counter factual trend of health care utilization under the fictitious world of no NHI, our subsequent difference-in-difference estimators tend to be downward biased for physician visits (without distinction between any visit and conditional visits) and any visit, and upward biased for conditional visits. However, there are many reasons to suspect that the counter factual trend in health care utilization would be different from the trend prior to NHI. For example, the sample individuals were of course older after NHI than before NHI, and we do not know how age affects health care utilization by individuals with differing health status. The previously uninsured are on average unhealthier than the previously insured in terms of a typical ADL score (Table 2.4). It could be that the previously uninsured may have increased the probability of any visit faster than the previously insured even without changes in health insurance status as they aged. Among a subset of the elderly who visited physicians at least once in the past month, the trends of conditional visits by the previously insured and uninsured may have begun to converge as they aged.

5. Estimation Strategies and Results

5.1 Two-Part Model

Following much of the literature, we consider a two-part model of health care utilization decisions. The first part determines whether individuals make visits to doctors, and the second part analyzes the number of visits conditional on any visit. In many developing countries, an important public health goal is to expand access to the health care system, as reflected in the any visit outcome. By conditioning on any visit, the second part is presumably less likely to be affected by factors that preclude access to physicians, such as strong aversion to western medicine, or very large visit costs. Individuals with no visits during the past month account for 59%, 50%, 38%, and 21% of respondents in 1989, 1993, 1996, and 1999, respectively.

Pohlmeier et al (1995) claims that the determinants of contacting physicians could be substantially different from the determinants of subsequent visits because different decision makers have the initiative in each process. The decision to contact physicians is up to patients while the decision to determine the intensity of treatment is often made by physicians. Unfortunately, our survey data do not reveal whether a particular visit to a physician is a first contact or a follow-up visit. Since the number of physician visits is measured with a one-month window in our data set, our measurement of any visit includes both first contacts and subsequent visits, while variation in conditional visits would come solely from subsequent visits. However, it is still possible for subsequent visits to be initial visits for new illnesses.

Supplier-induced demand for medical services has been a problem in Taiwan as well as in other countries where medical-service providers are paid on a fee-for-service (FFS) basis. Cheng (2003) includes a good description of payments to medical-service providers in Taiwan. (Quote from Cheng 2003)

> Like all open-ended health insurance systems relying on FFS payment of providers, Taiwan's NHI has experienced rapid increases in the volume of services, which, in turn, has led to charges of supplier-induced demand for services, many of which may not have been medically necessary. ... Experts in Taiwan appear to believe that the absolute level of fees paid by the NHI is too low and that many fees are considered to be below cost. In the absence of effective volume controls, providers' simplest response to low fees is to expand the volume of services they provide while reducing the resources going into each unit of service (for example, shortened visit length). The BNHI's[7] chief executive officer, Hong-Jen Chang, remarked that "Taiwan's doctors are well paid. But they work very, very hard to use volume to make up for the low fees." Ta-Fu Huang, chairman of the DoH's[8] Quality Commission, has written extensively about Taiwan's medical culture of the "three-minute patient visit" with physicians that is typical of doctors in Taiwan.

Given that physicians are motivated not only by the desire to provide high quality medical care, but also by economic incentives, health insurance could increase the number of conditional visits since physicians know that patients with health insurance face low out-of-pocket expenses and can afford frequent visits to physicians.

5.2 Fixed Effects Models

5.2.1 Physician visits without distinction between any visit and conditional visits

---
[7] The Bureau of National Health Insurance in Taiwan
[8] Department of Health in Taiwan

Before distinguishing between any visit and conditional visits, we estimate a reduced-form model of total number of visits, controlling for individual fixed-effects and using both linear and Poisson models. Because both 1996 and 1999 data are available, I can compare the short-term impact with the impact after four years. Additionally, we estimate the impact of health insurance using the two pre-NHI waves controlling for individual fixed effects, to examine the estimates when NHI is not used as the source of identification.[9] Individuals whose physician visits in relevant sample years differ by more than 20 in absolute value are not used in the regression analyses.[10] Excluding such individuals decreases the sample individuals for the linear model from 3413 to 3406 (years 1989 and 1993), from 3204 to 3198 (years 1993 and 1996), and from 3094 to 3084 (years 1996 and 1999).[11]

The linear model of physician visits (t=89 and 93, t=93 and 96, or t=93 and 99) is:

$$n_{it} = \beta_1 x_{it} + \beta_2 hins_{it} + \mu_i + \varepsilon_{it}. \tag{1}$$

The Poisson model of physician visits (t=89 and 93, t=93 and 96, or t=93 and 99) is:

$$E(n_{it} \mid x_{it}, hins_{it}, \mu_i) = \exp(\beta_1 x_{it} + \beta_2 hins_{it} + \mu_i), \tag{2}$$

where

n_{it} is the number of physician visits in the month immediately preceding the date of the interview,

$hins_{it} = 1$ if the individual had health insurance coverage in the last month,

---

9 Table 7 reports changes in health insurance status from 1989 to 1993. Almost all individuals who were insured in 1989 were covered in 1993 as well, and the majority of individuals who were not insured in 1989 remained uncovered in 1993. Between 1989 and 1993, several social health-insurance programs began: Farmer's Insurance (FI), which was established in 1985, was phased in by 1989; Parents of government employees gained coverage under Government Employees Dependents Insurance (GEDI) in 1989. Low-Income Household Insurance (LIHI) was launched in 1990. These new social insurance programs could affect some of 129 sample individuals who were newly insured in 1993, and thus the changes in health insurance status for the 129 individuals might be relatively exogenous. However, it is still possible that people self-selected into health insurance. For example, although health insurance coverage was provided automatically to members of farmers' associations under FI, membership in a farmers' association was voluntary. Alternatively, people could find a job in the government sector to give health-insurance coverage to elderly parents in poor health under GEDI.

10 Including such individuals does not change the estimation results meaningfully except for those using data from 1993 and 1999, in which the impact estimate of health insurance is substantially larger when excluding such individuals.

11 The sample individuals of the Poisson model are smaller than those of the linear model due to a technical reason. Fixed-effects Poisson models drop sample individuals with zero physician visits in both years.

$hins_{it} = 0$ if the individual did not have health insurance coverage in the last month,

$\mathbf{x}_{it}$ is a vector of time-varying demographic and financial variables,

μ_i captures individual-specific time-invariant unobservable characteristics, and

ε_{it} is remaining error,

and $(n_{it} \mid x_{it}, hins_{it}, \mu_i)$ follows a Poisson distribution in the Poisson model.

A vector of demographic and financial variables $\mathbf{x}_{it}$ includes potentially endogenous variables. For example, work status and financial status dummies such as owning real estate and owning savings/stocks could be positively correlated with unobserved health which, in turn, would be negatively correlated with physician visits. Thus, the estimated coefficients on these variables may be downward biased. The potential endogeneity problems of some variables (ADL scores and the dummy for living alone, in addition to the work dummy and the financial status dummies) are not particular to the current models, but potentially apply to all models in this study.

Because it is possible that the effect of health insurance on physician visits depends on an individual's health, we also estimate versions of (1) and (2) in which we interact the treatment variables $hins$ with dummy variables for health status based on ADL scores. The ADL score is constructed from responses to twelve typical ADL (Activities of Daily Living) and IADL (Instrumental Activities of Daily Living) questions using factor analysis techniques. Individuals are grouped into three groups according to their ADL scores. The upper, middle, and lower thirds are labeled as poor, average, and good health. From a policy perspective, it is important to examine whether health insurance benefits elderly adults in greater need more than elderly adults in better health.

Table 2.8 reports the regression results. Columns (1) through (4) rely for identification on within-individual changes before NHI, using data from 1989 and 1993. Columns (5) through (8) measure the impact of health insurance in the short run, using data from 1993 and 1996, while columns (9) through (12) measure the impact of health insurance four years after the introduction of NHI, using data from 1993 and 1999. For each set of years, we estimate four models: linear with no interactions, linear with interactions, Poisson with no interactions, and Poisson with interactions. As expected, standard errors are smaller in the Poisson model than in the linear model. Focusing on the

Poisson results, the estimated coefficients on health insurance are larger when the source of identification is not NHI. Health insurance coverage increased physician visits by 35% prior to NHI, while insurance coverage increased physician visits by 25% both one year and four years after the introduction of NHI. The impact of health insurance is the largest for those in the average health group when using NHI as the source of identification (36% increase one year after the introduction of NHI and 31% increase four years after the introduction of NHI). The impact is the largest for those with poor health using data prior to NHI (58% increase). These observations are consistent with the conjecture that changes in health insurance status prior to NHI are at least partly due to self-selection of those with greater need for health care acquiring health insurance.

5.2.2 Any visit to a physician

Any visit to a physician is modeled as a logit process, with t =89 and 93, t =93 and 96, or t=93 and 99:

$$y_{it}^{*} = \beta_{11}x_{it} + \beta_{12}hins_{it} + u_i + e_{1it} \tag{3}$$

y_{it}^{*} is the latent propensity to visit a doctor and is unobserved. y_{it} is an indicator variable for whether the individual saw a doctor at least once in the past month. We assume that $y_{it} = 1$ if $y_{it}^{*} > 0$, and $y_{it} = 0$ otherwise. u_i captures individual-specific time-invariant characteristics, and e_{1it} is the remaining error, which is assumed to follow a logistic distribution.

Columns (1) through (6) in Table 2.9 present the estimates of the average impact of health insurance on any visit, controlling for individual fixed-effects (FE) under the assumption that e_{1it} follows a logistic distribution (FE logit model). The FE logit models use maximum likelihood methods to estimate the coefficients, in which the conditional likelihood functions select only sample individuals who visited physicians in one of two years. The sample sizes for the FE logit models decrease for this reason. Columns (1) and (2) in Table 2.9 estimate the impact of heath insurance prior to NHI. Columns (3) and (4) look at the short-term impact of NHI, and columns (5) and (6) look at the impact after four years. Even-numbered columns allow for interactions between insurance status and health status. A significant increasing trend in the probability of visiting a physician is observed by looking at the coefficients on year dummies both between 1989 and 1993 and between 1993 and 1999 but not between 1993 and 1996. The ADL index and its square are significant at conventional levels in all columns except Column (2). Poor health, as captured by the ADL index, raises the probability of any visit but at a decreasing rate. Health insurance increases the probability of any visit in the short run after NHI while it has no significant impact four years after NHI, or prior to NHI. Further, the short-run impact is driven by those with average health, not by those with poor health. It is possible that newly available health insurance induced the elderly with better and moderate health to visit physicians for check-ups or the treatment of mild illnesses that was postponed in anticipation of NHI.[12] In contrast, the impact of health insurance prior to NHI and four years after the policy change was the largest for those with poor health, although the estimates are not statistically significant at conventional levels. Evaluating at the means of the independent variables and assuming that the individual fixed effects are zero, the short-term impact of health insurance (0.465) is to raise the probability of any visit by only 0.4 percentage points. However, one must be cautious in evaluating the magnitude, because the assumption of zero individual fixed effects is arbitrary. The same coefficient estimate could be interpreted to raise the probability of any visit by as much as 11.6 percentage points if individual fixed effects are large.[13]

5.2.3 Physician visits conditional on any visit

Fixed-effects models are feasible for count dependent variables (Hausman et al 1984).[14] The fixed-effects Poisson model, to be estimated in this section, determines the number of physician visits conditional on any visit (conditional visits). The Poisson model for conditional visits (t=89 and 93, t=93 and 96, or t=93 and 99) can be expressed as:

(4)

---

[12] The NHI Law passed the Congress in July 1994.

[13] $\dfrac{\partial \mathrm{Prob}(y_{it}=1 \mid \mathbf{x}_{it}, \mathrm{hins}_{it}, u_i)}{\partial \mathrm{hins}_{it}}$ is largest when $\beta_{11}x_{it} + \beta_{12}\mathrm{hins}_{it} + u_i$ is approximately zero, where

$$\mathrm{Prob}(y_{it}=1 \mid \mathbf{x}_{it}, \mathrm{hins}_{it}, u_i) = \frac{\exp(\beta_{11}x_{it} + \beta_{12}\mathrm{hins}_{it} + u_i)}{1+\exp(\beta_{11}x_{it} + \beta_{12}\mathrm{hins}_{it} + u_i)}.$$

$$E(n_{it}-1|n_{it}\geq 1, x_{it}, hins_{it}, v_i)=\exp(\beta_{21}x_{it}+\beta_{22}hins_{it}+v_i)$$

where v_i is individual-specific time-invariant unobservable characteristics.

The dependent variable is the number of physician visits in the last month, not including the first physician visit. One is subtracted from the number of physician visits because this avoids truncation problems in the subsequent econometric analysis. Conditional visits is defined only for those with at least one physician visit in the last month. The FE coefficient estimator is consistent under the assumption that the conditional mean is correctly specified, even when the Poisson distribution assumption is incorrect (Wooldridge 1999). In other words, for consistency of the estimator, we do not need to assume that $(n_{it}-1|n_{it}\geq 1, x_{it}, hins_{it}, v_i)$ follows a Poisson distribution as long as $E(n_{it}-1|n_{it}\geq 1, x_{it}, hins_{it}, v_i)=\exp(\beta_{21}x_{it}+\beta_{22}hins_{it}+v_i)$ is correct. However, it is necessary to adjust standard errors to accommodate a deviation of the error distribution from the Poisson distribution. Thus, robust standard errors (Wooldridge 1999) are reported in Table 2.10.

Columns (1) and (2) in Table 2.10 report the results using data from 1989 and 1993. Columns (3) and (4) show the short-run impact of NHI using data from 1993 and 1996. Columns (5) and (6) look at the impact of NHI after four years using data from 1993 and 1999. Even-numbered columns allow for heterogeneous impacts of health insurance for individuals with different health statuses.

Robust standard errors are substantially larger than non-robust standard errors, avoiding over-rejection of the null hypotheses that the effects are zero. Some significant coefficient estimates may suffer from omitted variable or simultaneity problems. For example, a negative association between changes in work status and changes in conditional visits in columns (5) and (6) could be due to unobserved health: Poor health could deprive people of working ability and send them to physicians. Also, unobserved poor health may force people to sell real estate or to use up savings to cover living expenses, thus creating a spurious negative correlation between changes in ownership of real estate and/or savings and changes in conditional visits. Poor health measured by the ADL index increases conditional visits at a decreasing rate, although the effects are not

[14] Textbook exposition of count models is available in Greene (1997), Cameron et al (1998), and Wooldridge (2001)

always statistically significant at conventional significance levels. The trend for conditional visits was increasing between 1989 and 1993, decreasing between 1993 and 1996, and increasing between 1993 and 1999, although none of the coefficients on the year dummies are statistically significant.

The coefficient estimates on health insurance are much larger in columns (1) and (2) than in the other columns. This could be due to self-selection into health insurance programs. People desiring to see physicians could self-select into health insurance programs prior to NHI. Column (3) shows that the short-term impact of health insurance was not significantly different from zero. However, column (4) shows that health insurance reduced conditional visits by 13% for those in poor health, although the effect is not statistically significant. This could be due to congestion of medical facilities caused by the introduction of NHI. Frequent users without health insurance could have been forced to reduce conditional visits because of the pent-up demand for medical services from new users after the introduction of NHI. Column (5) shows that health insurance has a positive and statistically significant effect on conditional visits four years after the introduction of NHI. Health insurance increased conditional visits by 29%. Further, Column (6) shows that those with poorer health have a larger impact of health insurance. A 42% increase in conditional visits for individuals with poor health is statistically significant at the five percent significance level.

5.2.4 Discussion

Prior to NHI, self-selection into health insurance was likely to be common. Although there was virtually no private health-insurance market in Taiwan, it was still possible that people who needed health care services could find ways to gain health insurance coverage. For example, farmers could obtain access to health insurance under Farmer's Insurance by becoming members of a farmer's association. Parents of government employees became eligible for health insurance under Government Employees' Dependents Insurance in 1989, so people could get a job in the government sector to insure unhealthy parents. Since the values of health insurance are higher for frequent users, it is possible that only such users were motivated to become newly insured, yielding the finding of no significant insurance impact on any visit but a highly significant and large impact on conditional visits.

Since coverage under NHI is mandatory for all Taiwanese residents by law, self-selection into NHI should not be a problem. Even using NHI as the source of identification, however, the impact of health insurance is different depending on when the impact is measured. In the short run, health insurance has a positive effect on any visit, especially for those with average health, although it has no significant impact on conditional visits. Four years after the introduction of NHI, health insurance had no significant impact on any visit for all health groups while it had a positive and significant effect on conditional visits, especially for those with poor health. The observed pattern of the impact of health insurance is consistent with the following interpretation. The policy change generated a period of adjustment in the short run, creating a large increase in demand to see physicians from the previously uninsured with average health, and a decrease in conditional visits by both previously insured and uninsured individuals with poor health, possibly because of congestion of medical facilities. In addition to Figures 2.6 and 2.8, national statistics are consistent with these stories. Data on physician visits in Table 2.1 are consistent with people refraining from seeing physicians in 1995, creating pent-up demand that surfaced in 1996. Table 2.11 presents statistics in the supply of medical facilities and personnel in Taiwan from 1988 to 2001. Both medical facilities and personnel expanded during the period. However, the total number of medical facilities increased at a slower rate in 1995 and 1996 than in previous years, and the number of physicians per 10,000 population did not increase in 1995 and 1996, making it plausible that medical facilities were crowded in 1996 due to pent-up demand that emerged in that year. Figure 2.10 plots the number of outpatient visits per physician from 1992 to 2001. The number peaks in 1993 and 1996, but based on our results first visits (that are time-intensive in comparison with follow-up visits) probably constituted a larger proportion in 1996 than in 1993. In total, it is likely that the short-term impact of health insurance was influenced by short-run adjustments in the demand for health care.

Focusing on the regression results after four years of the introduction of NHI, the impact of health insurance on any visit is not significantly different from zero for all health groups, while the impact of health insurance on conditional visits is positive and significant at the ten percent level. Further, poorer health makes the impact on conditional visits larger. As pointed out earlier, it is unlikely that frequent users are those with obstacles to access physicians. The significant and insignificant impact of health insurance on conditional visits and any visit respectively could imply that preferences and/or non-monetary costs (time and transportation costs) play a critical role in shaping the impact of health insurance. Unfortunately, detailed data on non-monetary costs of seeing physicians are not available in the survey data, so it is not possible to examine the effect of non-monetary costs on physician visits. Even if such data did exist, it might be difficult to control for fixed-effects because there are unlikely to be many exogenous within-individual changes in non-monetary costs (See Buchmueller et al 2004 for the effect of hospital closures on the use of some medical services in Los Angeles County).

5.3 Random Effects Model with Endogenous Health Insurance

5.3.1 Econometric model

Hausman et al (1984) introduced random-effects (RE) models for count dependent variables. However, I use a slightly different model here since I want to take into account the endogenous nature of health insurance coverage. The any-visit and conditional-visit equations in the following simultaneous-equation system are the same as before, except that we now include time-invariant variables as controls. In addition, a separate selection equation models health insurance coverage. Only pre-NHI data are used to estimate the system using the random-effects approach under the following assumptions about the error structure.

As before, we use the two-part decision making model, and the first part models whether or not an individual made any visit in the last month (RE probit model) with $t = 89, 93$.

$$y_{it}^{*} = \beta_{11} x_{it} + \beta_{12} hins_{it} + u_i + e_{1it} \tag{5}$$

where $y_{it} = 1$ if $y_{it}^{*} > 0$, and $y_{it} = 0$ otherwise.

The second part uses the RE negative binomial model for conditional visits with $t = 89, 93$.[15]

(6)

---

[15] A detailed description of the RE negative binomial model (conditional-visit equation) in the system is in the Appendix 2.1.

$$E(n_{it}-1\,|\,n_{it}\geq 1, x_{it}, hins_{it}, v_i)=\exp(\beta_{21}x_{it}+\beta_{22}hins_{it}+v_i)$$

The selection equation models whether an individual had health insurance coverage in the last month (RE probit model) with $t=89,93$.

$$hins^*_{it}=\beta_{31}x_{it}+\beta_{32}spaccess_i+w_i+e_{3it} \tag{7}$$

where $hins_{it}=1$ if $hins^*_{it}>0$, and $hins_{it}=0$ otherwise.

$$\begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_u & \sigma_{uv} & \sigma_{uw} \\ \sigma_{uv} & \sigma^2_v & \sigma_{vw} \\ \sigma_{uw} & \sigma_{vw} & \sigma^2_w \end{pmatrix}\right)$$

$e_{1it}\sim$ iid $N(0,0.78)$
$e_{3it}\sim$ iid $N(0,0.78)$
$(n_{it}-1\,|\,n_{it}\geq 1, x_{it}, hins, v_i)\sim$ Negative Binomial

where

$y_{it}=1$ if the individual visited physicians once or more in the last month

$y_{it}=0$ if the individual did not visit a physician in the last month

n_{it}: number of physician visits in the last month

$hins_{it}=1$ if the individual had health insurance coverage in the last month

$hins_{it}=0$ if the individual did not have health insurance coverage in the last month

$spaccess_i=1$ if the spouse of the respondent had access to occupational health insurance

$spaccess_i=0$ if the spouse of the respondent did not have access to occupational health insurance

x_{it}: vector of demographic and financial variables (including time-invariant characteristics)

u_i, v_i, w_i: individual-specific time-invariant unobservable characteristics

Under the assumption that the system is correctly modeled, the coefficient on health insurance $hins$ captures the impact of health insurance for individuals whose health insurance statuses are influenced by changes in the binary identifying instrument $spaccess$.[16] Note that $spaccess_i$ is time-invariant because it is based on the spouse's main

[16] It is widely accepted in the economic literature that with a binary identifying instrument, the identification of effects comes from the mean difference in the dependent variable (conditional on other covariates) for those with the identifying instrument being zero and for others with the identifying instrument being one. Angrist et al (1995) proved this in the context of two-stage least squares estimation. However, I am unaware of a rigorous proof that is valid when the system is estimated using the maximum likelihood method.

occupation in his or her lifetime. Thus, the identification of health insurance hinges on cross-sectional variation in this system, because the sole identifying instrument $spaccess_i$ is time-invariant.

Several assumptions are necessary to validate the simultaneous-equation system. First, all independent variables except the health insurance dummy must be uncorrelated with the unobserved effects. Second, the conditional mean of each dependent variable must be correctly specified. Third, the unobserved time-invariant individual effects (u_i, v_i, w_i) conditional on observed characteristics are jointly normally distributed with zero means. Finally, the time-varying errors conditional on the independent variables and the unobserved time-invariant individual characteristics for the any-visit and selection equations, i.e. $(e_{1it} \mid x_{it}, \text{spaccess}_i, u_i)$ and $(e_{3it} \mid x_{it}, \text{spaccess}_i, w_i)$, are normally distributed with mean zero. The time-varying error conditional on the independent variables and the unobserved time-invariant individual characteristics for the conditional-visit equation, i.e. $(n_{it} - 1 \mid n_{it} \geq 1, x_{it}, \text{spaccess}_{it}, v_i)$, follows a particular type of the negative binomial distribution (See the Appendix 2.1 for a detailed explanation).

5.3.2 Identification issues

The identifying IV spaccess_i must be a strong determinant of health insurance status but should not be correlated with the omitted factors that determine physician visits. Before the introduction of NHI, some elderly individuals had access to occupational health insurance. Retired government employees and their spouses, and veterans and their spouses are such individuals. Also, all farmers, regardless of age, were eligible for farmer's health insurance. For these reasons, having a government, military, or agricultural job as one's main occupation should strongly predict health insurance coverage before NHI. However, one problem with using the respondent's main occupation as an instrument is that it is likely to be correlated with other individual characteristics such as attitudes toward doctors or unobserved health. Fortunately, the respondent's spouse's main occupation is also available in the data set. The spouses of government and military workers, regardless of their own occupation, were eligible for health insurance even after retirement. Spouses of farmers also needed to be farmers to be eligible for occupational health insurance, but there was no rigorous monitoring of hours worked in farming, so an individual whose spouse was a farmer could easily qualify for insurance even if their main occupation in life was not farming. Table 2.12 lists the group of occupations that I assign $\text{spaccess}_i = 1$. Own occupation is controlled for in all equations in the system (dummies for main occupation in agriculture, the military, and government).

Bound et al (1995) proposed that researchers check the quality of the IV estimate in two ways: the F statistic and partial R^2 on the excluded instruments in the first-stage regression. Since we estimate the system using a maximum likelihood method, we look at the t statistic on spaccess_i and the likelihood ratio test on the excluded IV in the selection equation. The selection equation in the simultaneous-equation system yields the t statistic on spaccess_i of 3.50 where the standard error is Huber-corrected. Further, the likelihood ratio test on the excluded IV produces a p-value of 0.0001572,[17] suggesting that spaccess_i is a strong determinant of health insurance coverage before NHI.

5.3.3 Estimation results

Column (1) in Table 2.13 shows the coefficient estimates when the system is estimated simultaneously using the maximum likelihood method. Standard errors are Huber-corrected. The same variables have quite different impacts on any visit versus conditional visits. One additional year of schooling has a negligible impact on any visit for all individuals in the sample (max years of schooling observed: 17), while an additional year of schooling has a strong negative impact on conditional visits for individuals with higher education. The probability of any visit is significantly smaller for males than for females while the number of conditional visits is larger for males than for females after controlling for other covariates. Working or not does not matter for conditional visits, while working individuals seem to be less likely to make any visit to a physician than those who are not working. Ever receiving pensions increases the

[17] The selection equations with and without spaccess_i in the independent variables are estimated independently from other equations. The log likelihoods with and without spaccess_i in the explanatory variables are -1666.2369 and -1673.3788, respectively. With one degree of freedom, $\Pr(\chi^2 > 2\times 7.1419) = 0.0001572$.

probability of any visit and reduces the number of conditional visits. Bad health (measured by ADL score) increases both the probability of any visit and conditional visits. The number of adult children (whether they live in the same household or not) has a strong positive impact on conditional visits, while its impact on any visit is negligible. The year dummy shows that the trend in the probability of any visit rose with time. The trend in conditional visits is also increasing but the effect is significant only at the ten percent level.

As expected, having a major occupation in farming, the military, or government significantly increases the probability of health insurance coverage in the selection equation. Also, spouse access to health insurance strongly predicts health insurance coverage. With the endogenous nature of health insurance coverage corrected, health insurance has a negligible impact on any visit and a large positive impact on conditional visits. In terms of the magnitude, the average impact of health insurance on the probability of any visit is at most 0.0047 (0.47 percentage point),[18] which is a very small effect. In contrast, health insurance increases conditional visits by 63%. The coefficient estimate is statistically significant at the one percent level.

Column (2) in Table 2.13 estimates the system without the selection equation. Standard errors are Huber-corrected. The coefficient estimates on the health insurance dummies change, with effects on any visit of 0.312 and on conditional visits of 0.416, both of which are statistically significant at the one percent level.

Column (1) in Table 2.13 shows that for the full model the estimated error correlation between u_i and w_i (0.394) is significant at the one percent level, while the estimated error correlation between v_i and w_i (-0.184) is statistically insignificant. The strong positive correlation between u_i and w_i implies that there are time-invariant unobservable characteristics that increase the probability of any visit as well as the probability of health insurance coverage. In contrast, conditional visits seem to be less affected by self-selection. This implies that before the introduction of NHI, the insured were more likely to be those who were motivated to make at least one physician visit than the uninsured, while among the subset of individuals who visited physicians at least once in the past month, the insured were no more likely to be those who were motivated to utilize medical care services than the uninsured. This could occur if a significant proportion of the uninsured with at least one physician visit were motivated to visit physicians due to unobserved poor health, so that if one restricts attention to those with at least one physician visit, there is no or even a negative correlation between health insurance coverage and the motivation to visit physicians.

A few caveats about the RE simultaneous-equation model deserve mention. The identification of the impact of health insurance crucially relies on the excluded IV spaccess. If spouses are more likely to choose occupations with health insurance when the partners are unhealthy, spaccess is not a valid IV.[19] Further, because of assortative matching, married couples could share some characteristics that are correlated with physician visits. Alternatively, daily food they share could similarly affect the health of husbands and wives. If such unobserved characteristics are correlated with spouse access to health insurance, spaccess is not qualified as a valid IV, either. Finally, the coefficient estimates on the RE simultaneous-equation system are influenced by the arbitrary assumptions about the distributions of the unobserved individual effects and the remaining errors.

The estimation results using the RE simultaneous-equation system are similar to the fixed-effects results using pre-NHI data. Both results yield no significant impact of health insurance on any visit and a significant and large impact of health insurance on conditional visits of 60-70%. The fixed-effects results four years after the introduction of NHI also show no significant impact of health insurance on any visit and a significant but smaller impact (29%) of health insurance on conditional visits. The impact of health insurance on conditional visits could be different before and after the introduction of NHI because of the government's effort to contain health care costs.[20]

---

[18] $\left|\frac{\partial}{\partial z_{it,k}}\Phi(x_{it}\beta/\sigma_{\upsilon})\right| = \left|\frac{\beta_k}{\sigma_{\mu}}\varphi(x_{it}\beta/\sigma_{\upsilon})\right| = \left|\frac{-0.012}{\sqrt{0.514^2+0.885^2}}\varphi(x_{it}\beta/\sigma_{\upsilon})\right| \le 0.012\times\frac{1}{\sqrt{2\pi}} \approx 0.0047$

[19] However, the dependent portion of Government Employees Insurance started in 1982 and Farmer's Insurance started in 1985, making it unlikely that spouses chose *main occupations in his or her lifetime* in government or farming because of health insurance coverage provided to the partners.

[20] On the demand side, NHI introduced a new co-payment scheme in August 1999, including 20% co-payment on outpatient drug expenses over NT$100 and additional co-payments on high frequent visits and rehabilitation. On the supply side, to discourage induced-demand for medical services, a systematic financial monitoring mechanism was set up in July 1999 and a series of cost control measures were

6. Conclusions

In this chapter, I study the impact of health insurance on physician visits by the elderly in Taiwan. The existing literature on NHI in Taiwan measures the impact of health insurance immediately after the introduction of NHI, which reflects only short-term effects. This study finds evidence that physician visits in 1996 (one year after the introduction of NHI) were influenced by the pent-up demand for medical services and congestion of medical facilities. Exogenous policy changes are a popular source of identification in economics, because they are less endogenous in nature. However, a policy change can generate different results in the short, medium, and long terms.

Using pre-NHI data, the fixed-effects and IV methods yield estimates that are consistent with each other. Both find that the impact of health insurance on any visit is not significantly different from zero while the impact of health insurance on conditional visits is statistically significant and large (60-70%). The identification strategy using the natural experiment finds that after four years, NHI increased conditional visits by 29% and did not increase any visit. Given potential difficulties in the IV approach and possible endogenous changes in health insurance coverage before NHI, the identification based on the natural experiment should be more reliable. However, aggregate demand and supply changes affected by NHI complicate interpretation of the effects of NHI, and could partly explain the difference in the impact of health insurance on conditional visits. The different identification strategies also estimate the impact of health insurance for different subpopulations, which could produce different results if the impact of health insurance is heterogeneous.[21]

implemented (www.nhi.gov.tw/00english/e_03res_1.htm, accessed on September 6, 2004). The timing of these policy changes is consistent with a fall in outpatient medical utilization in 2000 as presented in Table 1 and Figure 2.10. However, since the 1999 survey interviews finished by June for the vast majority of the respondents, the government's effort to contain health care costs described here is unlikely to affect the number of physician visits reported by the respondents for 1999.

[21] The IV approach measures the impact of health insurance for the subpopulation whose health insurance coverage is affected by the spouse's main occupation. This subpopulation would include many housewives whose husbands (used to) work in farming, the military, or government. The fixed-effects approach using pre-NHI data measures the impact of health insurance for the subpopulation who changed health insurance coverage between 1989 and 1993. The subpopulation would include farmers, parents of government employees, and low-income households. The identification strategy based on the natural experiment measures the impact of health insurance for the subpopulation whose insurance coverage is affected by NHI. This subpopulation would include former workers in the private sector and their spouses.

Table 2.1: Trend in Medical Utilization in Taiwan

Year	# Outpatient	% Change	# Inpatient	% Change	# Emergency	% Change
1992	76,672,710	-	2,284,584	-	3,962,075	-
1993	81,583,740	6.41%	2,245,818	-1.70%	4,246,045	7.17%
1994	82,431,260	1.04%	2,391,242	6.48%	4,349,512	2.44%
1995	79,404,371	-3.67%	2,412,720	0.90%	4,664,209	7.24%
1996	86,134,506	8.48%	2,546,210	5.53%	4,992,277	7.03%
1997	89,109,624	3.45%	2,586,296	1.57%	5,257,705	5.32%
1998	93,550,483	4.98%	2,689,003	3.97%	5,459,637	3.84%
1999	96,703,254	3.37%	2,732,881	1.63%	5,883,886	7.77%
2000	96,074,268	-0.65%	2,823,800	3.33%	6,184,031	5.10%
2001	99,779,162	3.86%	2,922,513	3.50%	6,199,674	0.25%

(Source) www.stat.gov.tw (accessed on September 6, 2004)

Table 2.2: Health Related Indicators for Selected Countries
(1990 unless otherwise stated)

Country	Per capita GNP (1990 US$)	Share of GDP by agriculture	Life expectancy at birth (years)	Share population age 65 or older	c) Share population insured by health insurance	d) Share medical care to total hh consumption	Health expenditures as % of GDP	e) Doctors per 1000 population	f) Hospital beds per 1000 population
Low-income	*w) 350*	*w) 31%*	*w) 62*	*w) 4.5%*	-	-	-	-	-
India	350	31%	59	4.4%	5%	3%	6.0%	0.41	0.7
Kenya	370	b) 27%	59	2.8%	10%	3%	4.3%	0.14	1.7
Indonesia	570	22%	62	3.9%	13%	2%	2.0%	0.14	0.7
Middle-income	*w) 2220*	*w) 12%*	*w) 66*	*w) 6.1%*	-	-	-	-	-
Philippines	730	b) 22%	64	3.3%	38%	2%	2.0%	0.12	1.3
Dominican Rep.	830	b) 17%	67	3.4%	6%	8%	3.7%	1.08	2.0
Ecuador	980	b) 13%	66	3.6%	9%	5%	4.1%	1.04	1.7
Paraguay	1110	b) 28%	67	3.5%	18%	2%	2.8%	0.62	1.0
Columbia	1260	17%	69	4.0%	15%	7%	4.0%	0.87	1.5
Turkey	1630	18%	67	4.3%	58%	4%	4.0%	0.74	2.1
Panama	1830	b) 10%	73	4.7%	50%	8%	-	-	-
Costa Rica	1900	b) 16%	75	4.2%	82%	7%	-	-	-
Korea, Rep. Of	5400	b) 9%	71	5.5%	90%	5%	6.6%	0.73	3.0
Taiwan	**8111**	**4%**	**74**	**6.2%**	**47%**	**5%**	**4.2%**	**1.10**	**4.4**
High-income	*w) 19590*	-	*w) 77*	*w) 12.9%*	-	-	-	-	-
Netherlands	17320	b) 4%	77	13.2%	100%	11%	7.9%	2.43	5.9
France	19490	b) 4%	77	13.7%	100%	13%	8.9%	2.89	9.3
Germany	a) 22320	a) b) 2%	76	15.0%	75%	a) 13%	8.0%	2.73	8.7
Japan	25430	b) 3%	79	11.9%	100%	10%	6.5%	1.64	15.9

w) weighted average by population;
a) Data refer to the Federal Republic of Germany before unification.
b) GDP and its components are at purchaser values.
c) includes only social health insurance, not count private health insurance.
d) Data refer to either 1980 or 1985.

e) Each value refers to one particular but not specified year between 1988 and 1992.
f) Each value refers to one particular but not specified year between 1985 and 1990.
(Sources) www.stat.gov.tw and www.doh.gov.tw/english/statistics/Welcome.html for Taiwan (accessed on September 6, 2004)
World Development Report 1992 and *World Development Report 1993*, The World Bank, for other countries

Table 2.3: Comparison of Age and Gender Distribution between Completed Sample and Population

(Unit %)

	1989				1993			
	Female		Male		Female		Male	
Age Group	Population	Sample	Population	Sample	Population	Sample	Population	Sample
60-64	15.4	13.6	21.8	23.0	-	-	-	-
65-69	11.6	12.2	15.3	16.3	17.7	17.0	24.2	26.4
70-74	7.9	8.5	9.3	9.4	12.6	12.7	15.6	16.2
75-79	5.6	5.4	5.6	5.4	8.0	7.9	8.6	8.1
80-84	2.9	2.3	2.3	1.9	4.8	4.5	4.3	4.1
85+	1.5	0.9	0.9	1.1	2.5	1.7	1.7	1.5
total	44.9	42.9	55.1	57.1	45.5	43.7	54.5	56.3
	1996				1999			
	Female		Male		Female		Male	
Age Group	Population	Sample	Population	Sample	Population	Sample	Population	Sample
60-64	-	-	-	-	-	-	-	-
65-69	-	-	-	-	-	-	-	-
70-74	20.9	21.3	27.2	29.0	20.4	18.4	26.7	27.1
75-79	13.5	13.4	15.3	14.6	13.6	14.8	15.9	16.8
80-84	7.7	7.6	7.6	7.6	7.3	7.4	7.3	7.2
85+	4.6	3.5	3.2	3.1	4.9	4.3	3.9	4.0
total	46.7	45.7	53.3	54.3	46.3	44.9	53.7	55.1

Source for Population Distribution: Statistical Yearbook of the Republic of China 2002

Table 2.4: Mean Comparison of Demographic and Financial Characteristics between the Previously Insured and the Previously Uninsured

Health insurance status in 1993	Insured	Uninsured	Mean diff. test p-value
Age (1996)	73.471	74.752	0.000***
	(5.376)	(5.695)	
Male (1996)	0.595	0.435	0.000***
	(0.491)	(0.496)	
Years of Schooling (1996)	4.241	2.820	0.000***
	(4.634)	(3.881)	
Married (1996)	0.586	0.402	0.000***
	(0.493)	(0.491)	
Work (1996)	0.136	0.088	0.001***
	(0.343)	(0.284)	
Own Real Estate (1996)	0.600	0.428	0.000***
	(0.490)	(0.495)	
Own Savings and/or Stocks (1996)	0.365	0.289	0.001***
	(0.482)	(0.454)	
Own Other Assets such as	0.019	0.015	0.498
Businesses and Jewelry (1996)	(0.136)	(0.121)	
Any Evidence of Receiving	0.741	0.451	0.000***
Pension Payments in Any Year (1996)	(0.438)	(0.498)	
Self Reported Health Status (1996)	2.948	3.006	0.287
1 to 5; 1: Excellent, 5: Poor	(1.087)	(1.054)	
Self Reported Health Compared to Others of	1.986	2.017	0.371
Similar Ages (1996) 1 to 3; 1: Better, 3: Worse	(0.696)	(0.664)	
Activity of Daily Living (ADL) Score (1996)	0.175	0.322	0.012**
Larger score means poorer health.	(1.127)	(1.222)	
Live Alone (1996)	0.285	0.200	0.000***
	(0.452)	(0.400)	
Number of Adult Children (1996)	4.627	4.365	0.004***
	(2.262)	(2.199)	
Rural Resident (1996)	0.619	0.387	0.000***
	(0.486)	(0.488)	

1) Standard deviations are in parentheses.
2) Male, Years of Schooling, and Any Evidence of Receiving Pension Payments are time- invariant, but due to panel attrition, the means differ across years.
3) The number of observations for the previously insured ranges from 1760 to 1915, depending on a specific variable. The number of observations for the previously uninsured ranges from 472 to 545, depending on a specific variable.

Table 2.5: Changes in Physician Visits by Health Insurance Status prior to NHI

	Health-insurance status prior to NHI	1989	1993	1996	1999
Physician Visits	The insured	1.181	1.531	1.966	2.449
	N=1018	(2.401)	(2.358)	(2.377)	(2.871)
	The uninsured	0.996	1.223	2.055	2.399
	N=238	(1.988)	(2.513)	(2.113)	(2.923)
Any Visit	The insured	0.422	0.565	0.687	0.817
	N=1042	(0.494)	(0.496)	(0.464)	(0.387)
	The uninsured	0.363	0.412	0.694	0.759
	N=245	(0.482)	(0.493)	(0.462)	(0.428)
Conditional Visits	The insured	1.845	1.985	1.760	2.085
	N=200	(3.189)	(2.978)	(2.421)	(2.816)
	The uninsured	1.111	2.037	2.296	2.852
	N=27	(2.006)	(4.024)	(1.836)	(3.231)

Including Outliers

	Health-insurance status prior to NHI	1989	1993	1996	1999
Physician Visits	The insured	1.235	1.599	1.970	2.565
	N=1028	(2.644)	(2.688)	(2.371)	(3.320)
	The uninsured	0.996	1.343	2.054	2.397
	N=239	(1.984)	(3.123)	(2.109)	(2.917)
Conditional Visits	The insured	1.854	2.200	1.766	2.420
	N=205	(3.174)	(3.754)	(2.398)	(3.897)
	The uninsured	1.071	3.000	2.250	2.786
	N=28	(1.980)	(6.446)	(1.818)	(3.190)

1) The mean numbers of physician visits, any visit, and conditional visits are shown. Standard deviations are in parentheses.
2) The table includes only those who reported the numbers of physician visits in ALL the years.
3) The insured means those with health insurance in 1989, 1993, 1996, and 1999 while the uninsured means those without health insurance in 1989 and 1993 but with health insurance in 1996 and 1999.
4) The outliers are defined as those whose reports on physician visits in adjacent sample years differ more than 20 in absolute value. The outliers are not relevant to Any Visit.

Table 2.6: Mean Changes in Physician Visits between 1989 and 1993 by Health Insurance Status

	The insured	The uninsured	p-value	Including Outliers The insured	The uninsured	p-value
Physician Visits	0.310	0.198	0.442	0.337	0.258	0.627
	(0.074)	(0.125)		(0.082)	(0.139)	
	n=1767	n=475		n=1773	n=476	
Any Visit	0.127	0.021	0.001***			
	(0.015)	(0.028)		Same		
	n=1783	n=476				
Conditional Visits	-0.058	0.602	0.168	-0.066	0.940	0.0879*
	(0.191)	(0.436)		(0.205)	(0.547)	
	n=483	n=83		n=485	n=84	

1) Standard errors of the corresponding means are in parentheses.

Table 2.7: Changes in Health Insurance Status from 1989 to 1993

89 insurance status	93 insurance status Insured	Uninsured	Total
Insured	1994	84	2078
	95.96%	4.04%	100.00%
Uninsured	129	574	703
	18.35%	81.65%	100.00%
Total	2123	658	2781
	76.34%	23.66%	100.00%

1) Frequencies and row relative frequencies are shown.

Table 2.8: Fixed-Effects Linear Regression Results and Fixed-Effects Poisson Regression Results for Physician Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years used	89&93	89&93	89&93	89&93	93&96	93&96	93&96	93&96	93&99	93&99	93&99	93&99
Linear or Poisson FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age Squared/100	-0.134	-0.142	-0.091	-0.096*	0.093	0.076	0.095	0.076	-0.226	-0.226	-0.078*	-0.076
	(0.133)	(0.133)	(0.058)	(0.058)	(0.194)	(0.194)	(0.079)	(0.079)	(0.140)	(0.141)	(0.047)	(0.047)
Married	0.084	0.101	0.063	0.065	-0.115	-0.12	-0.145*	-0.146*	-0.079	-0.096	-0.064	-0.073
	(0.208)	(0.207)	(0.086)	(0.086)	(0.205)	(0.205)	(0.083)	(0.083)	(0.237)	(0.238)	(0.076)	(0.076)
Work	0.005	0.035	-0.004	0.023	-0.262	-0.262	-0.164**	-0.177**	-0.205	-0.18	-0.190***	-0.180**
	(0.138)	(0.138)	(0.058)	(0.059)	(0.174)	(0.174)	(0.069)	(0.069)	(0.216)	(0.216)	(0.071)	(0.072)
Own Real Estate	-0.047	-0.046	-0.066	-0.068	-0.232*	-0.229	-0.166***	-0.170***	-0.103	-0.109	-0.045	-0.049
	(0.139)	(0.138)	(0.059)	(0.059)	(0.140)	(0.140)	(0.054)	(0.054)	(0.151)	(0.152)	(0.047)	(0.047)
Own Saving/Stocks	-0.110	-0.109	-0.074	-0.074	-0.084	-0.095	-0.070*	-0.076*	-0.098	-0.098	-0.087*	-0.089*
	(0.108)	(0.108)	(0.046)	(0.046)	(0.106)	(0.106)	(0.041)	(0.041)	(0.141)	(0.141)	(0.045)	(0.046)
Own Other Assets	0.115	0.129	0.091	0.117	-0.241	-0.261	-0.104	-0.123	-0.125	-0.13	-0.015	-0.002
	(0.247)	(0.247)	(0.104)	(0.105)	(0.240)	(0.240)	(0.091)	(0.091)	(0.339)	(0.339)	(0.107)	(0.108)
ADL Score	0.673***	0.129	0.460***	0.131	1.032***	1.031***	0.463***	0.508***	1.248***	1.105***	0.552***	0.522***
	(0.186)	(0.236)	(0.073)	(0.097)	(0.180)	(0.239)	(0.063)	(0.084)	(0.231)	(0.340)	(0.069)	(0.102)
ADL Score Squared	-0.138***	-0.044	-0.099***	-0.042**	-0.183***	-0.180***	-0.086***	-0.092***	-0.242***	-0.213***	-0.111***	-0.103***
	(0.041)	(0.048)	(0.016)	(0.020)	(0.039)	(0.047)	(0.014)	(0.016)	(0.054)	(0.069)	(0.016)	(0.021)
Live Alone	-0.080	-0.072	-0.097	-0.082	0.042	0.032	0.051	0.042	0.140	0.129	0.076	0.074
	(0.153)	(0.152)	(0.062)	(0.063)	(0.157)	(0.157)	(0.058)	(0.058)	(0.183)	(0.183)	(0.057)	(0.057)
# Adult Children	0.073	0.066	0.054*	0.052*	0.039	0.038	0.035	0.032	0.062	0.061	0.072**	0.068**
	(0.078)	(0.078)	(0.030)	(0.030)	(0.079)	(0.079)	(0.029)	(0.029)	(0.096)	(0.096)	(0.031)	(0.031)

Table 2.8 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years used	89&93	89&93	89&93	89&93	93&96	93&96	93&96	93&96	93&99	93&99	93&99	93&99
Linear/Poisson FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE	Lin. FE	Lin. FE	Poi. FE	Poi. FE
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Own Insurance	0.306		0.352***		0.353**		0.252***		0.296		0.248***	
(HINS)	(0.229)		(0.109)		(0.149)		(0.061)		(0.200)		(0.065)	
HINS * (Better Health)		0.103		0.195*		0.242		0.187**		0.116		0.170**
		(0.243)		(0.118)		(0.181)		(0.077)		(0.244)		(0.081)
HINS*(Average Health)		0.131		0.217*		0.501***		0.362***		0.397*		0.306***
		(0.246)		(0.115)		(0.169)		(0.069)		(0.223)		(0.073)
HINS * (Poor Health)		0.819***		0.577***		0.319*		0.197***		0.386		0.261***
		(0.265)		(0.118)		(0.192)		(0.071)		(0.278)		(0.083)
Year Dummy 1989	-1.011	-1.051	-0.729**	-0.746**								
	(0.742)	(0.741)	(0.322)	(0.323)								
Year Dummy 1996					-0.06	0.02	-0.196	-0.112				
					(0.842)	(0.843)	(0.340)	(0.342)				
Year Dummy 1999									2.681**	2.650**	1.049**	1.027**
									(1.236)	(1.238)	(0.412)	(0.412)
Constant	7.507	8.159			-3.972	-3.061			11.998*	12.052*		
	(6.907)	(6.891)			(9.860)	(9.881)			(7.112)	(7.126)		
Observations	5806	5806	3176	3176	5423	5423	3478	3478	4753	4753	2954	2954
# Individuals	3406	3406	1588	1588	3198	3198	1739	1739	3084	3084	1477	1477
Log likelihood			-2710.17	-2696.00			-2842.67	-2835.07			-2407.36	-2404.55

1) Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%

2) Individuals whose physician visits in relevant two years differ more than 20 in absolute value are not included.

Table 2.9: Fixed-Effects Logit Regression Results for Any Visit

	(1)	(2)	(3)	(4)	(5)	(6)
Years used	89&93	89&93	93&96	93&96	93&99	93&99
Insurance interacted w/ health	No	Yes	No	Yes	No	Yes
Age Squared/100	-0.191	-0.190	-0.107	-0.132	-0.464***	-0.472***
	(0.144)	(0.144)	(0.222)	(0.223)	(0.146)	(0.147)
Married	-0.031	-0.007	-0.054	-0.044	-0.291	-0.314
	(0.215)	(0.215)	(0.229)	(0.231)	(0.279)	(0.279)
Work	-0.144	-0.131	-0.123	-0.143	-0.026	0.000
	(0.150)	(0.151)	(0.204)	(0.206)	(0.244)	(0.245)
Own Real Estate	0.028	0.035	-0.250	-0.261	-0.066	-0.086
	(0.141)	(0.141)	(0.166)	(0.167)	(0.180)	(0.181)
Own Saving/Stocks	0.072	0.071	-0.053	-0.085	-0.019	-0.011
	(0.114)	(0.115)	(0.120)	(0.121)	(0.161)	(0.162)
Own Other Assets	0.068	0.065	-0.368	-0.459	-0.182	-0.187
	(0.284)	(0.284)	(0.281)	(0.288)	(0.393)	(0.392)
ADL Score	0.523***	0.250	0.774***	0.846***	1.109***	0.681*
	(0.202)	(0.263)	(0.206)	(0.281)	(0.274)	(0.392)
ADL Score Squared	-0.112***	-0.066	-0.147***	-0.156***	-0.238***	-0.164**
	(0.042)	(0.051)	(0.043)	(0.053)	(0.060)	(0.076)
Live Alone	0.236	0.238	-0.049	-0.035	-0.069	-0.089
	(0.165)	(0.165)	(0.177)	(0.178)	(0.208)	(0.209)
Number Adult Children	-0.022	-0.019	0.036	0.033	-0.114	-0.116
	(0.082)	(0.082)	(0.095)	(0.097)	(0.105)	(0.105)
Own Health Insurance	-0.088		0.465***		0.099	
(HINS)	(0.240)		(0.174)		(0.222)	
HINS * (Better Health)		-0.182		0.298		-0.199
		(0.258)		(0.211)		(0.280)
HINS * (Average Health)		-0.224		0.754***		0.105
		(0.262)		(0.199)		(0.249)
HINS * (Poor Health)		0.144		0.361		0.393
		(0.280)		(0.228)		(0.311)
Year Dummy 1989	-1.478*	-1.477*				
	(0.802)	(0.804)				
Year Dummy 1996			1.074	1.197		
			(0.968)	(0.973)		
Year Dummy 1999					5.476***	5.521***
					(1.309)	(1.315)
Observations	2010	2010	1858	1858	1518	1518
Number of Individuals	1005	1005	929	929	759	759
Log likelihood	-662.21	-660.46	-566.33	-560.83	-346.85	-345.33

1) Standard errors in parentheses

2) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.10: Fixed-Effects Poisson Regression Results for Conditional Visits

	(1)	(2)	(3)	(4)	(5)	(6)
Year used	89&93	89&93	93&96	93&96	93&99	93&99
Insurance interacted w/ health	No	Yes	No	Yes	No	Yes
Age Squared/100	-0.116	-0.114	0.227	0.215	-0.083	-0.085
	(0.193)	(0.186)	(0.193)	(0.196)	(0.118)	(0.119)
Married	0.081	0.059	-0.113	-0.098	0.036	0.036
	(0.325)	(0.335)	(0.215)	(0.218)	(0.205)	(0.207)
Work	0.039	0.071	0.054	0.029	-0.368*	-0.362*
	(0.161)	(0.157)	(0.145)	(0.140)	(0.198)	(0.199)
Own Real Estate	0.011	0.011	-0.236*	-0.246**	-0.122	-0.127
	(0.187)	(0.192)	(0.121)	(0.122)	(0.122)	(0.122)
Own Saving/Stocks	-0.418***	-0.420***	-0.046	-0.039	-0.087	-0.093
	(0.145)	(0.141)	(0.105)	(0.105)	(0.131)	(0.129)
Own Other Assets	0.040	0.115	-0.087	-0.124	-0.111	-0.102
	(0.328)	(0.350)	(0.276)	(0.275)	(0.281)	(0.278)
ADL Score	0.320	-0.107	0.171	0.344*	0.635***	0.440*
	(0.223)	(0.268)	(0.150)	(0.206)	(0.190)	(0.259)
ADL Score Squared	-0.045	0.033	-0.018	-0.047	-0.103**	-0.068
	(0.049)	(0.054)	(0.031)	(0.039)	(0.046)	(0.056)
Live Alone	-0.219	-0.210	0.074	0.065	0.092	0.101
	(0.183)	(0.171)	(0.137)	(0.138)	(0.129)	(0.129)
Number Adult Children	-0.017	-0.018	0.055	0.056	0.078	0.079
	(0.120)	(0.123)	(0.109)	(0.104)	(0.078)	(0.077)
Own Health Insurance	0.714**		-0.016		0.289*	
(HINS)	(0.353)		(0.151)		(0.160)	
HINS * (Better Health)		0.425		0.061		0.184
		(0.379)		(0.195)		(0.202)
HINS * (Average Health)		0.624*		0.116		0.231
		(0.379)		(0.164)		(0.181)
HINS * (Poor Health)		0.967**		-0.126		0.415**
		(0.378)		(0.173)		(0.189)
Year Dummy 1989	-0.692	-0.667				
	(1.072)	(1.037)				
Year Dummy 1996			-0.877	-0.827		
			(0.835)	(0.849)		
Year Dummy 1999					0.685	0.696
					(1.038)	(1.046)
Observations	1002	1002	1416	1416	1270	1270
Number of Individuals	501	501	708	708	635	635
Log likelihood	-959.02	-949.66	-1162.83	-1159.39	-1099.9	-1098.25

1) Robust standard errors in parentheses

2) * significant at 10%; ** significant at 5%; *** significant at 1%

3) Individuals whose physician visits in relevant two years differ more than 20 in absolute value are not included.

Table 2.11: Supply-Side Indicators of Medical Services

Year	Medical care facilities a) b)			Beds a) b)		No. health personnel per 10,000 population b)		
	Total	per medical care facility		Total	No. beds per	Total c)	Physicians	Nurses
	(numbers)	Area served (km^2)	Population served (persons)	(number)	10,000 population	(persons)	(persons)	(persons)
1988	12215 (-)	2.95 (-)	1629 (-)	88572 (-)	44.50 (-)	41.7 (-)	10.3 (-)	17.0 (-)
1989	12267 (0.4%)	2.93 (-0.7%)	1639 (0.6%)	86693 (-2.1%)	43.11 (-3.1%)	42.6 (2.2%)	10.4 (1.0%)	17.6 (3.5%)
1990	12902 (5.2%)	2.79 (-4.8%)	1578 (-3.7%)	89151 (2.8%)	43.80 (1.6%)	44.8 (5.2%)	11.0 (5.8%)	18.9 (7.4%)
1991	13661 (5.9%)	2.64 (-5.4%)	1505 (-4.6%)	92785 (4.1%)	45.14 (3.1%)	47.2 (5.4%)	11.5 (4.5%)	20.3 (7.4%)
1992	14468 (5.9%)	2.49 (-5.7%)	1434 (-4.7%)	96084 (3.6%)	46.30 (2.6%)	49.6 (5.1%)	12.0 (4.3%)	22.1 (8.9%)
1993	15062 (4.1%)	2.40 (-3.6%)	1394 (-2.8%)	100570 (4.7%)	47.90 (3.5%)	52.2 (5.2%)	12.5 (4.2%)	24.0 (8.6%)
1994	15752 (4.6%)	2.30 (-4.2%)	1344 (-3.6%)	103733 (3.2%)	48.98 (2.3%)	53.9 (3.3%)	12.9 (3.2%)	25.4 (5.8%)
1995	16109 (2.3%)	2.25 (-2.2%)	1326 (-1.3%)	112379 (8.3%)	52.62 (7.4%)	55.4 (2.8%)	12.9 (0.0%)	26.6 (4.7%)
1996	16645 (3.3%)	2.17 (-3.6%)	1293 (-2.5%)	114923 (2.3%)	53.39 (1.5%)	57.5 (3.8%)	12.9 (0.0%)	28.6 (7.5%)
1997	17398 (4.5%)	2.08 (-4.2%)	1250 (-3.3%)	121483 (5.7%)	55.87 (4.7%)	63.4 (10.3%)	13.4 (3.9%)	32.0 (11.9%)
1998	17731 (1.9%)	2.04 (-1.9%)	1237 (-1.0%)	124564 (2.5%)	56.80 (1.7%)	65.7 (3.6%)	14.0 (4.5%)	32.5 (1.6%)
1999	17770 (0.2%)	2.04 (0.0%)	1243 (0.5%)	122937 (-1.3%)	55.65 (-2.0%)	69.0 (5.0%)	14.4 (2.9%)	34.2 (5.2%)
2000	18082 (1.8%)	2.00 (-2.0%)	1232 (-0.9%)	126476 (2.9%)	56.78 (2.0%)	71.5 (3.6%)	15.0 (4.2%)	35.5 (3.8%)
2001	18265 (1.0%)	1.98 (-1.0%)	1227 (-0.4%)	127676 (1.0%)	56.98 (0.4%)	74.0 (3.5%)	15.4 (2.7%)	36.9 (3.9%)

1) Percentage increases from previous years are in parentheses.

a) Number of public and private hospitals and clinics

b) Beginning in 1994, data include the Taiwan-Fukien area. (The Taiwan-Fukien area contains 0.3% of the total population in 2001.)

c) Medical personnel total includes physicians, dentists, pharmacists and assistants, nurses, medical laboratory technicians and assistants, medical radio technologists and assistants, and midwives. Beginning in 1995, dietitians are included. Beginning in 1997, physician therapists and physician therapist assistants are included.

(Source) Statistical Yearbook of the Republic of China 2002

Table 2.12: Occupations with Access to Health Insurance to Spouses before 1995

Code	Occupation
03	teachers of middle school or above (including professors)
04	teachers of elementary school, kindergarten, and child-care centers
05	lawyers and judges (prosecutors, and chief judges and heads of courts)
15	heads of state-owned enterprises (not staff but give orders)
16	military officers (highly ranked officers, Junior Lieutenant or above)
25	secretaries (government or state-owned enterprises)
26	accountants (government or state-owned enterprises)
27	staff members other than secretaries and accountants (committee members or above) (government or state-owned enterprises)
28	other staff members with lower-level positions (such as temporary staff, (typewriting staff, and file managers) (government or state-owned enterprises)
48	other technicians or artisans, or "skilled workers" (government or state-owned enterprises)
55	military officers (senior sergeant or below)
65	policemen, firemen
81	farmers (growing flowers or vegetables, or raising pigs)
82	farm owners
83	landlords
84	farm workers (employed)
85	farmers (unspecified)
86	fishermen (boat owners or captains)
87	fishermen (employed to fish)
88	mine workers
89	forest workers

Table 2.13: Regression Results of Simultaneous-Equation System

Any Visit Probit Equation	(1)	(2)
Correct Endogenous *hins*?	Yes	No
Age / 10_1	1.120 *	1.176 *
	(0.618)	(0.614)
Age Squared / 100_1	-0.080 *	-0.083 *
	(0.043)	(0.043)
Male_1	-0.107 *	-0.108 *
	(0.064)	(0.063)
Years of Schooling_1	0.017	0.011
	(0.013)	(0.013)
Years of Schooling Squared_1	-0.001	-0.000
	(0.001)	(0.001)
Married_1	0.007	-0.008
	(0.047)	(0.047)
Work_1	-0.137 ***	-0.140 ***
	(0.048)	(0.047)
Own Real Estate_1	0.082 *	0.075 *
	(0.044)	(0.044)
Own Saving/Stocks_1	-0.018	-0.018
	(0.042)	(0.042)
Own Other Assets_1	-0.035	-0.029
	(0.109)	(0.109)
Ever Received Pensions_1	0.127 ***	0.084 *
	(0.048)	(0.046)
ADL Score_1	0.507 ***	0.499 ***
	(0.074)	(0.074)
ADL Score Squared_1	-0.096 ***	-0.095 ***
	(0.018)	(0.018)
Height (in meters)_1	0.570 *	0.497
	(0.336)	(0.330)
Live Alone_1	0.023	0.017
	(0.047)	(0.046)
No. Adult Children_1	-0.007	-0.009
	(0.010)	(0.010)
Own Health Insurance_1	-0.012	0.312 ***
	(0.104)	(0.053)
Major Job in Farming_1	0.041	-0.043
	(0.059)	(0.054)
Major Job in Military_1	-0.011	-0.09
	(0.086)	(0.082)
Major Job in Government_1	0.159 **	0.098
	(0.080)	(0.077)
Rural Resident_1	0.007	-0.021
	(0.049)	(0.047)
Year Dummy 1989_1	-0.283 ***	-0.271 ***
	(0.040)	(0.040)
Constant_1	-5.064 **	-5.288 **
	(2.282)	(2.269)

Table 2.13 (Continued)

Conditional Visit Equation	(1)	(2)
Correct Endogenous *hins*?	Yes	No
c (Look at Appendix 2.1.)	0.352 ***	0.351 ***
	(0.104)	(0.104)
Age / 10_2	0.935	0.922
	(1.061)	(1.063)
Age Squared / 100_2	-0.073	-0.073
	(0.073)	(0.073)
Male_2	0.260 ***	0.258 ***
	(0.098)	(0.098)
Years of Schooling_2	0.026	0.029
	(0.022)	(0.022)
Years of Schooling Squared_2	-0.003 *	-0.003 *
	(0.002)	(0.002)
Married_2	-0.130 *	-0.120
	(0.078)	(0.077)
Work_2	-0.073	-0.071
	(0.087)	(0.087)
Own Real Estate_2	-0.045	-0.040
	(0.074)	(0.074)
Own Saving/Stocks_2	-0.298 ***	-0.298 ***
	(0.073)	(0.073)
Own Other Assets_2	0.298	0.298
	(0.201)	(0.203)
Ever Received Pensions_2	-0.123	-0.099
	(0.079)	(0.074)
ADL Score_2	0.532 ***	0.534 ***
	(0.110)	(0.110)
ADL Score Squared_2	-0.094 ***	-0.095 ***
	(0.025)	(0.026)
Height (in meters)_2	-1.005 *	-0.955 *
	(0.521)	(0.517)
Live Alone_2	0.005	0.006
	(0.073)	(0.073)
No. Adult Children_2	0.044 ***	0.045 ***
	(0.016)	(0.016)
Own Health Insurance_2	0.626 ***	0.416 ***
	(0.228)	(0.092)
Major Job in Farming_2	-0.088	-0.037
	(0.102)	(0.088)
Major Job in Military_2	-0.302 **	-0.255 *
	(0.145)	(0.141)
Major Job in Government_2	-0.04	-0.004
	(0.141)	(0.136)
Rural Resident_2	0.136 *	0.152 *
	(0.081)	(0.080)
Year Dummy 1989_2	-0.119 *	-0.126 *
	(0.071)	(0.070)
Constant_2	-2.229	-2.161
	(4.004)	(4.007)

Table 2.13 (Continued)

Insurance Probit Equation	(1)	(2)
Correct Endogenous *hins*?	Yes	No
Age / 10_3	-1.525	
	(1.867)	
Age Squared / 100_3	0.096	
	(0.130)	
Male_3	0.301	
	(0.284)	
Years of Schooling_3	0.200 ***	
	(0.061)	
Years of Schooling Squared_3	-0.004	
	(0.005)	
Married_3	0.513 ***	
	(0.166)	
Work_3	0.224	
	(0.152)	
Own Real Estate_3	0.244 *	
	(0.143)	
Own Saving/Stocks_3	0.202	
	(0.128)	
Own Other Assets_3	-0.371	
	(0.294)	
Ever Received Pensions_3	1.416 ***	
	(0.217)	
ADL Score_3	0.054	
	(0.193)	
ADL Score Squared_3	-0.023	
	(0.043)	
Height (in meters)_3	2.420 *	
	(1.357)	
Live Alone_3	0.072	
	(0.158)	
No. Adult Children_3	0.082 **	
	(0.040)	
Major Job in Farming_3	2.446 ***	
	(0.293)	
Major Job in Military_3	2.675 ***	
	(0.433)	
Major Job in Government_3	1.974 ***	
	(0.367)	
Rural Resident_3	1.068 ***	
	(0.207)	
Spouse Access to Insurance	0.816 ***	
	(0.233)	
Year Dummy 1989_3	-0.440 ***	
	(0.110)	
Constant_3	-0.224	
	(7.018)	

Table 2.13 (Continued)

Error Estimation	(1)	(2)
Correct Endogenous *hins*?	Yes	No
sigma_u	0.514 ***	0.491 ***
	(0.048)	(0.048)
sigma_v	0.683 ***	0.680 ***
	(0.045)	(0.044)
sigma_w	3.062 ***	
	(0.249)	
rho_uv	0.493 ***	0.558 ***
	(0.155)	(0.150)
rho_uw	0.394 ***	
	(0.104)	
rho_vw	-0.184	
	(0.184)	
sigma_e1	0.885	0.885
sigma_e3	0.885	
Sample Size	4645	4645
	2055	2055
	4645	-
Log Likelihood	-8333.49	-6678.20

1) Asymptotic standard errors in parentheses;
2) Significance: '*'=10%; '**'=5%; '***'=1%.
3) All standard errors are Huber-corrected.
4) Individuals whose physician visits in 1989 and 1993 differ by more than 20 in absolute value are not included in estimating the conditional- visit equation in the system.

Figure 2.1: National Trend of Outpatient Visits

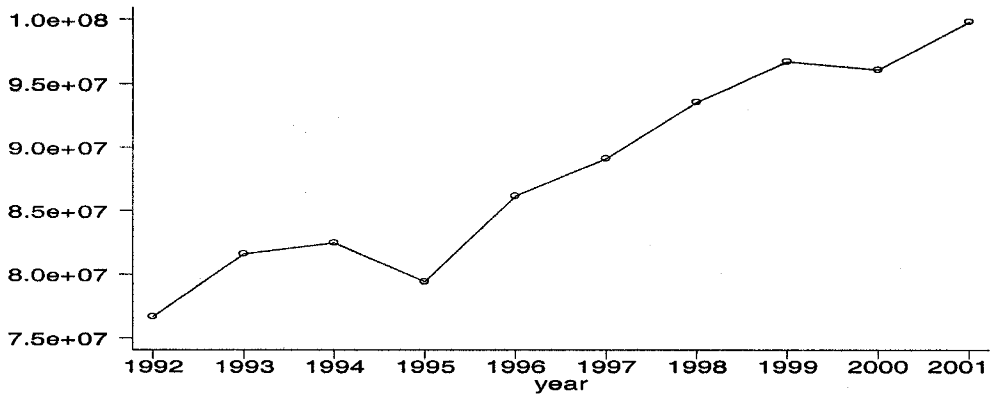


Figure 2.2: National Trend of Inpatient Visits

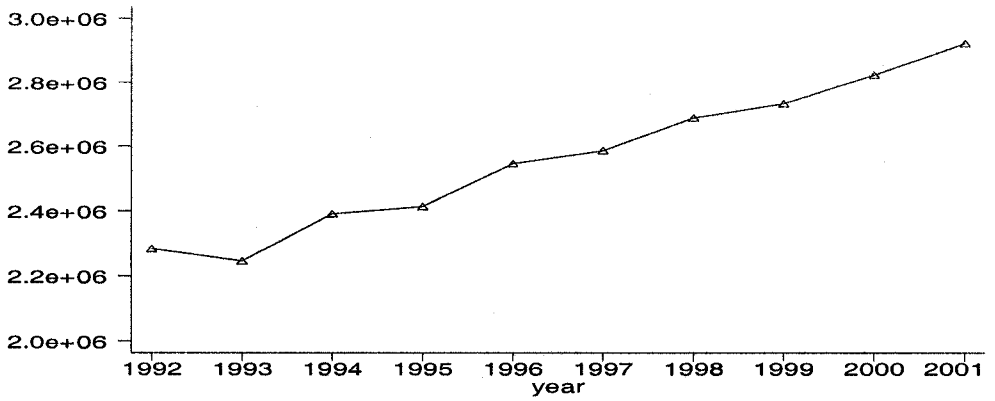
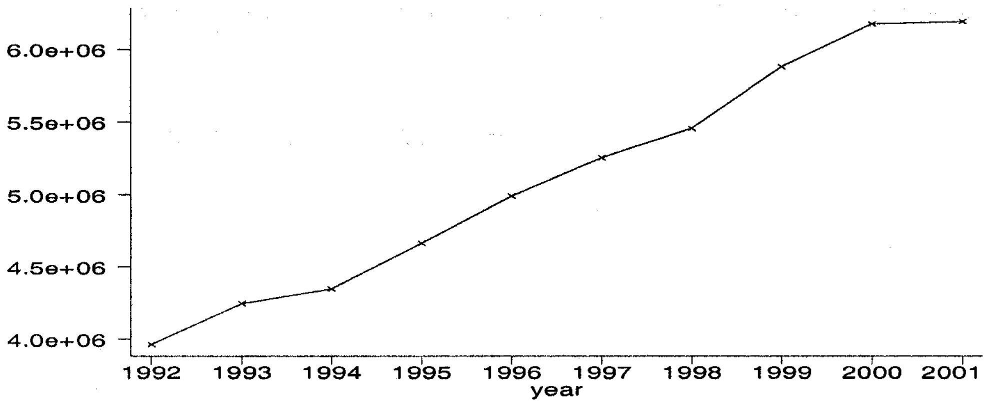


Figure 2.3: National Trend of Emergency Visits



(Source) www.stat.gov.tw (accessed on September 6, 2004)

Figure 2.4: Sample Trend of Physician Visits

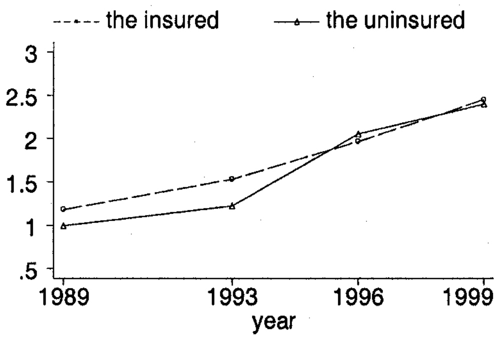


Figure 2.5: Sample Difference in Physician Visits

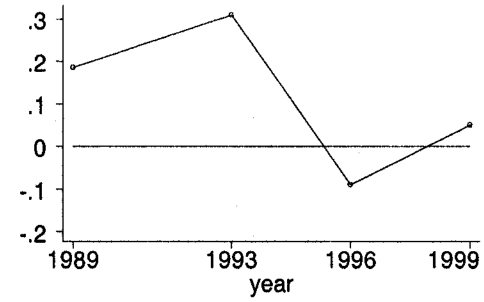


Figure 2.6: Sample Trend of Any Visit

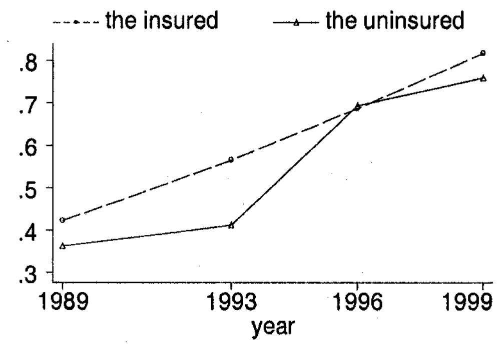


Figure 2.7: Sample Difference in Any Visit

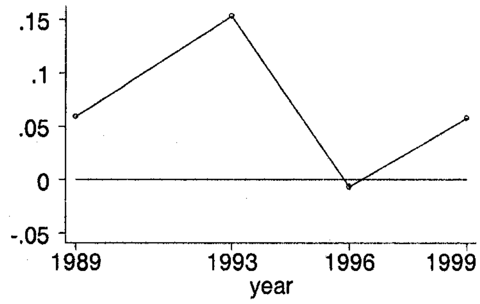


Figure 2.8: Sample Trend of Conditional Visits

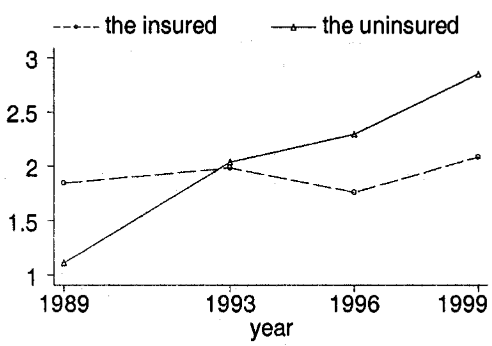


Figure 2.9: Sample Difference in Conditional Visits

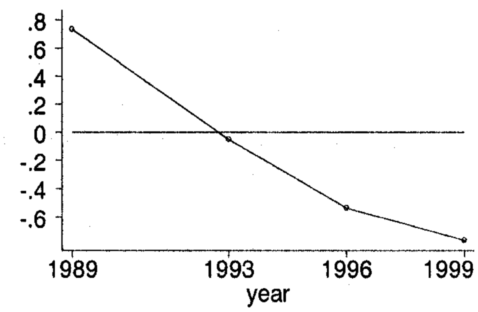
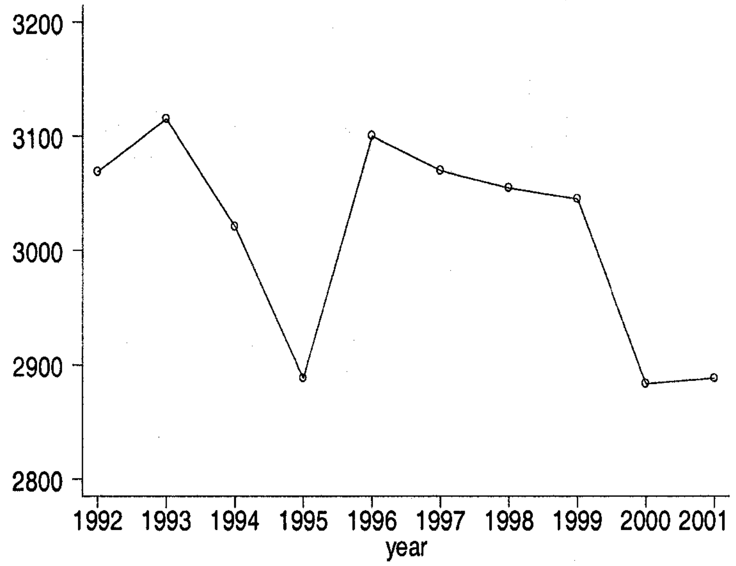


Figure 2.10: National Trend of Outpatient Visits per Physician



(Figure 2.10 Source)
www.stat.gov.tw (accessed on September 6, 2004) and author's calculation

Appendix 2.1 Explanation of the Negative Binomial Model (Conditional-Visit Equation) in the RE Simultaneous-Equation System

In order to allow for pairwise correlations between the time-invariant individual-specific unobservable characteristics ($u_i, v_i,$ and w_i), the joint normality of $u_i, v_i,$ and w_i is assumed in the simultaneous-equation system. The drawback of this approach is that the closed form of the conditional probability function is not available. (Thus, a closed form of the likelihood function does not exist.) This is the reason why Hausman et al (1984) assume that the unobserved time-invariant error is Gamma or Beta distributed in deriving the panel count data models. However, even with the joint normality assumption, the simultaneous-equation system is estimable numerically. I use aML (Lillard and Panis 2000) to estimate the simultaneous-equation system. The aML writes the probability density of a negative binomial distribution as follows:

$$f(Y \mid A,p) = \frac{\Gamma(A+Y)}{\Gamma(A)\Gamma(Y+1)}(1-p)^A p^Y \qquad (1)$$

where Y is the count outcome given scale A and probability p, and $\Gamma(\cdot)$ denotes the Gamma function.[22]

The negative binomial probability density (1) implies

$$E(Y \mid A,p) = \frac{A}{p} - A = A\left(\frac{1-p}{p}\right)$$

and

$$\mathrm{Var}(Y \mid A,p) = \frac{A(1-p)}{p^2}.$$

I parameterized A (scale) and p (probability) as follows.

$$A = \exp(c) \text{ where c is constant} \qquad (2)$$

and

$$p_{it} = \frac{1}{1+\exp(\beta_{21}x_{it} + \beta_{22}hins_{it} + v_i)} \qquad (3)$$

Using the parameterizations (2) and (3),

[22] A typical example of a negative binomial distribution is as follows: You need to have A successes. The probability of success is 1-p, which is constant throughout trials. How many failures do you experience before you have A successes? The number of failures Y before the A[th] success follows a negative binomial distribution.

$$E(Y_{it} \mid A, p_{it}) = A\left(\frac{1-p_{it}}{p_{it}}\right) = \exp(c + \beta_{21}\mathbf{x}_{it} + \beta_{22}\text{hins}_{it} + v_i)$$
$$= \alpha_i^* \lambda_{it} = E(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}, v_i)$$

and

$$\text{Var}(Y_{it} \mid A, p) = \frac{A(1-p)}{p^2}$$
$$= \exp(c + \beta_{21}\mathbf{x}_{it} + \beta_{22}\text{hins}_{it} + v_i)[1 + \exp(c + \beta_{21}\mathbf{x}_{it} + \beta_{22}\text{hins}_{it} + v_i)]$$
$$= \alpha_i^* \lambda_{it}(1 + \alpha_i \lambda_{it}) = \text{Var}(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}, v_i)$$

where $\alpha_i^* = \exp(c + v_i)$, $\alpha_i = \exp(v_i)$, and $\lambda_{it} = \exp(\beta_{21}\mathbf{x}_{it} + \beta_{22}\text{hins}_{it})$. Thus, the variance to mean ratio of $n_{it} - 1$ conditional on $n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}$, and unobserved v_i is given by $(1 + \alpha_i \lambda_{it})$. Since $E(v_i) = 0$ and $\text{Var}(v_i) > 0$,

$$E(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}, v_i) = E(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it})$$

and

$$\text{Var}(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}, v_i) < \text{Var}(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it}),$$

implying that the variance to mean ratio conditional only on observed characteristics is larger than one (overdispersion). However, other properties of the variance to mean ratio conditional only on observed characteristics are not clear since the closed form of the conditional probability function $f(n_{it} - 1 \mid n_{it} \geq 1, \mathbf{x}_{it}, \text{hins}_{it})$ is not available.

References

Angrist, Joshua, and Guido Imbens. (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90: 431-442.

Bound, John, David A. Jaeger, and Regina M. Baker. (1995). "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90: 443-450.

Buchmueller, Thomas C., Mireille Jacobson, and Cheryl Wold. (2004). "How Far to the Hospital? The Effect of Hospital Closures on Access to Care," mimeo, UC Irvine and NBER.

Cameron, Colin, and Pravin Trivedi. (1998). *Regression analysis of count data*, Cambridge University Press.

Cheng, Shou-Hsia, and Tung-Liang Chiang. (1997). "The Effect of Universal Health Insurance on Health Care Utilization in Taiwan," *Journal of the American Medical Association*, 278: 89-93.

Cheng, Tsung-Mei. (2003). "Taiwan's new national health insurance program: Genesis and experience so far," *Health Affairs*, 22 (3): 61-76.

Chiang, Tung-liang. (1997). "Taiwan's 1995 Health Care Reform," *Health Policy*, 39: 225-239.

Chou, S.Y., Jin-Tan Liu, and James K. Hammitt. (2002). "Health Insurance and Households' Precautionary Behaviors – An Unusual Natural Experiment," NBER Working Paper 9394.

Chou, Y.J. and Douglas Staiger. (2001). "Health insurance and female labor supply in Taiwan," *Journal of Health Economics*, 20: 187-211.

Directorate-General of Budget, Accounting and Statistics Executive Yuan, Republic of China. (2002). *Statistical Yearbook of the Republic of China 2002.*

Greene, William. (1997). *Econometric Analysis*, Third Edition, Prentice Hall.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. (1984). "Econometric Models for Count Data with an Application to the Patents-R & D Relationship," *Econometrica*, 52: 909-938.

Imbens, Guido, and Joshua Angrist. (1994). "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62: 467-475.

Lillard, Lee, and Constantijn Panis. (2000). *aML Release 1User's Guide and Reference Manual*, Econware, Inc.

Manning, Willard, Joseph Newhouse, Naihua Duan, Emmett Keeler, and Arleen Leibowitz. (1987). "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review*, 77(3): 251-277.

McWilliams, J. Michael, Alan M. Zaslavsky, Ellen Meara, and John Z. Ayanian. (2003). "Impact of Medicare Coverage on Basic Clinical Services for Previously Uninsured Adults," *Journal of the American Medical Association*, 290 (6): 757-764.

Peabody, JW, Yu, JC-I, Wang Y-R, and Bickel SR. (1995). "Health System Reform in the Republic of China: Formulating Policy in a Market-Based Health System," *Journal of the American Medical Association*, 273: 777-781.

Pohlmeier, W., and V. Ulrich. (1995). "An Econometric Model of the Two-Part Decision-Making Process in the Demand for Health Care," *Journal of Human Resources*, 30: 339-361.

Taiwan Provincial Institute of Family Planning, and Population Studies Center, University of Michigan. (1989). *1989 Survey of Health and Living Status of the Elderly in Taiwan.*

————————. (1997). *1996 Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan.*

Tsai, Wei-Der, and Jin-Tan Liu. (1998). "Health Insurance and the Demand for Ambulatory Care by the Elderly in Taiwan," Discussion Paper No. 9803, Institute of Economics, Academic Sinica.

Wooldridge, Jeffrey. (1999). "Distribution-Free Estimation of Some Nonlinear Panel Data Models," *Journal of Econometrics*, 90: 77-97.

Wooldridge, Jeffrey. (2001). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

World Bank. (1992). *World Development Report 1992*, Oxford University Press.

World Bank. (1993). *World Development Report 1993*, Oxford University Press.

Chapter 3

The Effect of Water Accessibility on Child Health in China

Abstract

Using panel data from China, the effect of access to clean water on child health is measured by incorporating child-specific fixed effects for the first time. Although fixed effects control for unobserved time-invariant characteristics that are correlated with both water access and child health, two sources of potential bias remain. First, as long as households have control over access to clean water, their decisions to invest in access to clean water may depend on unobserved factors that are correlated with changes in child health. For example, unobserved changes in wealth may increase both the ability to pay for water access and the demand for child health. This chapter addresses this problem by choosing a subset of sample children whose water access was likely affected by external projects rather than households' own demand. However, even when access to clean water is provided by government or NGO investments in water projects, which are exogenous to individual household decisions, the placements of water projects may not be random. Such investments could be correlated with other community dynamics such as investments in other health-related projects. To deal with unobserved community dynamics that are potentially correlated with both changes in water access and changes in child health, we make use of community fixed effects. Addressing the unobserved dynamic confounding factors, this chapter finds that having access to clean water within the yard of one's house improves child health.

1. Introduction

The World Health Organization (WHO) estimates that, today, 1.6 million people per year die due to unsafe water and lack of basic sanitation. In addition, water-associated diseases, in particular malaria and filariasis, present another heavy burden, with more than one million deaths due to malaria alone. Children are considered especially vulnerable to insufficient quality and quantity of drinking water. About 400 children below age five die per hour in the developing world from waterborne diarrheal diseases. While daily per capita consumption of two liters is the generally accepted value for a person weighing 60kg,[1] infants and children consume more per unit weight than adults (Gadgil 1998). A sufficient amount of water is important for basic human needs not only for drinking but also for other purposes such as cooking and sanitation. In comparison with the range of water use in industrial countries (350 to 850 liters per person per day; De Zuane 1997), the definitions of acceptable water quantities for rural areas in the majority of developing countries (15-50 liters per person per day; WHO 1996) seem ungenerous. Given that some people in developing countries consume less than 15-50 liters per day, the quantity as well as quality of water use could be an important determinant of health.

WHO (2000) reports that an estimated 1.1 billion do not have access to safe drinking water on the globe. The Millennium Development Goal (MDG) agreed at the United Nations Millennium Summit in 2000 set the target of halving the proportion of people without access to an improved source of drinking water by 2015. In accordance with this agreement, countries across the world have pledged to provide 1.5 billion people with access to improved drinking water by 2015.

The attainment of the MDG is in line with China's own national development plans, which have been strongly committed to provide clean drinking water to its residents. The Chinese government announced that by the end of 2020, every rural family in China should have clean drinking water. Nowadays, more than 300 million rural residents throughout China still lack clean drinking water. Over the past five years, more than 14 million rural families throughout 27 provinces have gained access to drinking water, with more than 800,000 new water processing facilities going into operation. China earmarked a record 18 billion yuan (US$ 2.1 billion) for rural drinking water supply during the 10th Five-Year Plan (2001-2005). The funds mostly come from local revenues and national debt (China Daily, November 29, 2004).

[1] The actual water intake, however, varies considerably from individual to individual, and also according to climate, physical activity, and culture. Water need increases sharply as ambient temperature exceeds 25 degrees in centigrade, primary to make up for moisture loss through perspiration (Gadgil 1998).

Measuring the effect of water accessibility on child health is an important policy question. Water-supply projects have been popular in China as well as in many developing countries. Although much human labor and monetary resources have been devoted to water projects, the effect of clean water on child heath (a principal outcome of interest) is far from conclusive.

Previous research finds a positive relationship between access to clean water and child health (Merrick 1985, Cebu Study Team 1991, Thomas et al 1992, Lee et al 1997, and Jalan et al 2001). None of these studies, however, exploits changes in health within individual children before and after changes in water accessibility using panel data. In this chapter, child-specific fixed effects available using panel data control for time-invariant characteristics that are likely to be correlated with both water accessibility and child health. Such characteristics include parents' knowledge about health and community characteristics affecting child health.

Nonetheless, even with child fixed effects, measuring the effect of water accessibility on child health using observational data is plagued with potential problems. First, unobserved changes in wealth may increase both water access and child health through better nutrition and health care, creating upward bias. Alternatively, households could react to bad child health by investing in better access to clean drinking water, creating downward bias. More generally, as long as households have control over access to clean water, many unobserved factors affecting the household could be correlated with both changes in water access and changes in child health. In contrast, if access to clean water is mostly determined by government or NGO investments in water projects, access to clean water would be largely exogenous to households' own demand.

However, even if changes in access to clean water are mostly determined by governments or NGOs, the placements of projects may not be random across villages (Pitt et al 1999, Molyneaux et al 2000, Frankenberg et al 2001). For example, local governments or NGOs could give priority to areas where other health-related infrastructure is deteriorating quickly, leading to downward bias. Alternatively, investments in water access could be either positively or negatively correlated with other health-related investments. On the one hand, investments in water access may crowd out other health-related projects due to budget constraints, creating downward bias. On the other hand, if communities can afford investments in water access, they may also be able to afford other health-related projects, leading to upward bias. Note that with child fixed effects, project placements endogenous to time-invariant characteristics do not bias the estimated effect of water accessibility on child health. Rather, the identification issue is whether the dynamic characteristics of project sites are correlated with both changes in water access and changes in child health. This chapter addresses these problems by using the following two methods in addition to controlling for child fixed effects. First, to avoid a potential household-level dynamic correlation between changes in water access and unobserved changes in households' demand for water access, we use a subset of sample children whose water access was likely affected by external projects rather than households' own demand. Second, to address a potential community-level dynamic correlation between the placements of water projects and investments in other health-related projects, we make use of community fixed effects.

The effect of maternal schooling on child health or more generally home production has been discussed in the literature (Leibowitz 1974, Behrman et al 1987, Datcher-Loury 1988, Behrman et al 1989, Thomas et al 1991). Access to clean water and maternal education could be either complements or substitutes in producing child health (Thomas et al 1992). On the one hand, educated mothers may make better use of convenient access to clean water than less educated mothers by, for example, encouraging children to have better hygiene with the aid of easy access to clean water. On the other hand, educated mothers may be more skillful than less educated mothers in obtaining safe water (e.g. boil water) when easy access to clean water is unavailable. In this case, convenient access to clean water could benefit children of uneducated mothers more than children of educated mothers. Whether parental education and access to clean water are complements or substitutes has important policy implications. It might be that increasing parental education makes clean water projects more effective in promoting child health. Alternatively, clean water projects could decrease a gap in child health between educated and less educated households.

The rest of the chapter is organized as follows. Section 2 describes the data. It also provides summary statistics partly in order to make sure that there is enough within-child variation on the variables used in the econometric analysis. Econometric models and results are presented in Section 3. Section 4 concludes.

2. Data

Data from the second (1991) and third (1993) waves of the China Health and Nutrition Survey (CHNS) are used for the analyses.[2] The CHNS is one of the few datasets from developing countries that has information on child anthropometrics as well as household-level accessibility to clean water over time, making it possible to control for child fixed effects in examining the effect of water accessibility on child health.

Each wave of the CHNS consists of a household survey, individual surveys of health and nutrition, an elderly survey, an ever-married women survey, a community survey, and a health and family planning facility survey. The survey population is drawn from eight of China's thirty-one provinces, located throughout the country: Guangxi, Guizhou, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. A multistage, random cluster approach was used to construct the sample in each of the eight provinces. The 190 primary sampling units consisted of 32 urban neighborhoods, 30 suburban neighborhoods, 32 towns, and 96 villages. The household survey includes information on household income and assets (including how to obtain drinking water), as well as time allocation by household members.[3]

Table 3.1 describes key variables used for the econometric analyses in this chapter. We use the sample children whose ages are between two and fifteen (inclusive) in 1991. Body Mass Index (BMI)[4] for age standardized using healthy children in the United States as the reference population is used as the child health measure, following much of the development literature.[5] We do not use sample children whose BMI z scores in either 1991 or 1993 are three or larger in absolute value, because the heights and/or

---

[2] Complete data on child anthropometrics are not available in the first (1989) and fourth (1997) waves.
[3] Further information of the CHNS is available at http://www.cpc.unc.edu/projects/china (accessed on February 11, 2005).
[4] $\text{BMI} = \dfrac{\text{weight}}{\text{height}^2}$ where weight and height are measured in kilograms and meters, respectively.
[5] The formula and parameters to standardize BMI for age are available from the Center for Disease Control and Prevention of the US Department of Health and Human Services at http://www.cdc.gov/nchs/about/major/nhanes/growthcharts/datafiles.htm (accessed on February 11, 2005).

weights of those children are likely to be misreported (249 children of total 2915 children are excluded). Further, we do not include the sample children whose BMI z scores in either 1991 or 1993 are two or larger to exclude obese children (145 children of total 2666 children are excluded).

Water accessibility is measured by the dummy variable "near water," which indicates whether or not the household has a clean water source within the yard of one's house. The description of other variables is also included in Table 3.1.

Table 3.2 presents summary statistics for the sample children whose BMI z scores are between –3 and 2 in both sample years (which reduces the total number of children to 2521). It provides not only the cross-sectional means and standard deviations of key variables in this chapter but also the standard deviations of the within-child variations of those key variables. We see considerable within-child variations for the time-varying variables.

Since this chapter measures the effect of water accessibility on child health using child-specific fixed effects, within-child variation in water accessibility is of particular interest. As seen in Table 3.3, approximately 12% of 2516 sample children experienced changes in "near water" status between 1991 and 1993.

Table 3.4 presents the ratios of the sample children who decreased and increased their BMI z scores between the two sample years separately for the four types of sample children who experienced differing access to clean water between 1991 and 1993. Looking at all sample children, approximately 50% of children saw improved health. However, 60% of sample children who gained water access experienced health increases.

3. Econometric Analyses

3.1 Base-line Child-Specific Fixed-Effects Model

I estimate the following differenced linear child health equation.

$$\Delta Z_{it} = a + b\Delta W_{jt} + c\Delta Y_{jt} + \mathbf{f}\Delta \mathbf{V}_{kt} + \mathbf{g}\mathbf{D}_{d} + \mathbf{h}\mathbf{R}_{r} + \mathbf{q}\mathbf{M}_{m} + \Delta\varepsilon_{it} \tag{1}$$

where

subscript i indexes a child;

subscript t indexes time ($t = 1991, 1993$);

subscript j indexes a household;

subscript k indexes a community;

$\Delta Z_{it} = Z_{i,t} - Z_{i,t-1}$ is the differenced BMI z score for child i;

$\Delta W_{jt} = W_{j,t} - W_{j,t-1}$ is the differenced dummy indicating "near water" status;

$\Delta Y_{jt} = Y_{j,t} - Y_{j,t-1}$ is the differenced log deflated per-capita household income;

$\Delta V_{kt} = V_{k,t} - V_{k,t-1}$ is the differenced vector of community characteristics affecting child health;

$\Delta \varepsilon_{it}$ is the differenced time-varying error;

$\mathbf{D_d}$, $\mathbf{R_r}$, and $\mathbf{M_m}$ are control dummies explained below;

and

a, b, c, $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$, and $\mathbf{q}$ are the coefficients to be estimated.

$\mathbf{D_d}$ is a vector of demographic dummies, including a gender dummy (omitted category: girls) and age-group dummies[6] (omitted category: ages 13-15 in 1991). $\mathbf{R_r}$ is a vector of regional dummies, including provincial dummies[7] (omitted category: Jiangsu province) and a dummy for village residents (omitted category: non-village residents). $\mathbf{M_m}$ is a vector of interview-month dummies for both 1991 and 1993[8] (omitted category: for both 1991 and 1993, interviewed in October). These interview-month dummies control for seasonal changes in BMI z score. For example, children could be better fed in harvest months than in other months.

Standard errors in Tables 3.5, 3.7, 3.8, and Appendix 3.1 are robust to household-level clustering and heteroskedasticity. Column (1) in Table 3.5 reports the coefficient estimates of the base-line child-specific fixed-effects model. Having access to clean water within the yard of one's house ("near water") is positively correlated with child health. In

---

[6] Five age-group dummies are created based on age in 1991: 2-3 years old, 4-6 years old, 7-9 years old, 10-12 years old, and 13-15 years old.

[7] Eight provincial dummies are created: Liaoning, Henan, Shandong, Hubei, Hunan, Jiangsu, Guangxi, and Guizhou.

[8] Three interview-month dummies are created for each year: interviewed in September or before, interviewed in October, and interviewed in November or December.

terms of magnitude, "near water" is associated with a 0.178 increase in the child BMI z score, and it is statistically significant at the one percent significance level.

Changes in the BMI z score are systematically correlated with some demographic, regional, and interview-month dummies. Children with age six or younger in 1991, on average, experienced less positive changes in BMI z scores in comparison with the reference children (ages 13, 14, and 15 in 1991). Changes in the percentage of children (age 12 or younger) who received any immunization are the only community characteristic that is statistically significant at the conventional significance levels. Changes in the ratio of immunized children are negatively correlated with changes in the BMI z score, which is not intuitive. It could be that local governments gave priority to areas where immunization would be most effective: communities that experienced some negative shocks to child health.

Log deflated per-capita household income unexpectedly has a negative coefficient estimate although it is statistically insignificant. Because the contribution to household income by children is at most limited, I expected household income to be relatively exogenous in Equation (1). However, adolescent children may contribute to household income, so the coefficient estimate may be biased if changes in child health affect household labor supply. Adult labor supply could also be influenced by child health if sick children require greater care. Nevertheless, these stories do not explain a downward bias in estimating the effect of household income on child health. To explain the result, we need other stories. For example, adults could work longer to earn income to pay for medical expenses for sick children.

3.2 Addressing Potential Endogeneity of Household Income

Changes in household income are important in Equation (1) because it is plausible that increases in household income would increase both child health and water access. Without controlling for changes in household income, the estimated effect of access to clean water on child health would be upward biased. The effect of household income on child health is also of interest in itself.

We address the endogeneity concern described earlier in two ways. First, we restrict the sample to children aged ten or younger in 1991. It is quite safe to assume that the contributions to household income by children aged twelve or younger are minimal. Column (2) in Table 3.5 uses only children whose ages are ten or younger in 1991. The coefficient estimate for household income is still negative and statistically insignificant. The coefficient estimate for "near water" falls in its statistical significance to the ten percent level due to both a smaller coefficient estimate (0.152) and a larger standard error (0.085).

Another approach to address the potential correlation of unobserved changes in child characteristics with both changes in the BMI z score and changes in household income is to use instruments for changes in household income. Rainfall is an exogenous variable that affects agricultural income. Thus, rainfall variation is a good candidate to serve as an instrument for household income. We use monthly county-level rainfall data to construct instruments that capture variation in rainfall. Specifically, monthly rainfall data for the 58 sample counties are standardized using historic monthly rainfall data for the years 1961 to 1990, and the instruments are the number of standard deviations that monthly rainfall differs from historic monthly means (negative numbers if below the monthly averages).[9] As instruments, changes over time in the standardized amounts of rainfall in the following months of the current calendar year are used: March and September. First-stage results find that changes in rainfall in the other months did not have a significant effect on changes in household income.

We next address potential problems with using rainfall variation as instruments for household income. It is possible that rainfall could act as a productivity shock affecting the labor supply of individuals in agriculture, which, in turn, could influence BMI. This influence (through work effort) could be immediate or sequential. For instance, rainfall could change the amount of labor required in later stages of cultivation (e.g. low rainfall ruins the harvest, reducing required harvest labor, see also Fafchamps (1993) and Skoufias (1993)).

---

[9] Historic climate data collected from more than 250 climate stations all over China are publicly available (Two Long-Term Instrumental Climatic Data Bases of the People's Republic of China, compiled by the Chinese Academy of Sciences). The University of North Carolina (UNC) merged the CHNS counties with the climate data, using an interpolation algorithm called Inverse Distance Weighting (IDW). IDW assigns the weighted average of climate data to each county, where weights are the inverses of the distances to the county from a group of surrounding climate stations located within 300km from the target county.

To test whether the instruments are correlated with child labor supply, I regress changes in child farm work hours[10] on the set of the variables used for the first-stage regressions of changes in household income. The excluded instruments are jointly not significant, failing to reject the hypothesis that the instruments are uncorrelated with changes in child labor supply. Whenever rainfall variation is used as instruments for changes in household income, the tests of joint significance of the excluded instruments in the regressions of changes in farm work hours as well as changes in household income are reported in the bottom of Table 3.5.

Column (3) in Table 3.5 presents the coefficient estimates when changes in household income are instrumented by using changes in standardized rainfall. The coefficient estimate for household income becomes positive and highly significant. A ten percent increase in household income is associated with a 0.035 increase in the BMI z score. The coefficient for "near water" indicates a 0.204 increase in the BMI z score with better water access and is significant at the one percent significance level.

3.3 Addressing Potential Endogeneity of Accessibility to Clean Water

To deal with the endogeneity of changes in water access, we focus on households whose water accessibility was likely affected by external projects rather than households' own demand for access to clean water. In other words, I exploit the fact that changes in water access due to household-specific circumstances are isolated. In each community, I look at the number of sample households which gained or lost access to clean water between 1991 and 1993. Community here is defined as either city or suburban neighborhood in urban areas and town or village in rural areas. Table 3.6 shows the distribution of the number of sample households within communities for 1991 and 1993. Each community includes, on average, 19.0 sample households in 1991 and 18.5 sample households in 1993. Column (4) in Table 3.5 uses only sample communities which satisfy one of the following two conditions: (i) sample communities where more than fifteen percent of sample households changed access to clean water in one direction

[10] Farm work hours in the past year immediately preceding the date of the interview.

(either gained or lost access to clean water),[11] or (ii) sample communities where no sample households changed water access between 1991 and 1993. The control group under this identification strategy includes the following two groups of children: (i) children who experienced no changes in water access and lived in communities where more than 15% of sample households changed water access; and (ii) children who lived in communities where no sample households changed water access. The treatment group under this identification strategy is sample children who experienced changes in water access and who lived in communities where more than 15% of sample households changed water access. This identification strategy relies on the assumption that it would have been on average no difference in changes in the BMI z score between the treatment group and the control group if it had not been for water-supply projects in the communities. Since we use the arbitrary criterion of fifteen percent, the sensitivity of the results to other criteria (5%, 10%, 20%, and 25%) is examined later. Households could lose access to "near water" exogenously when droughts dry up wells, for example. The sample size decreases to 1883 children with the fifteen-percent restriction. The number of sample children who gained, lost, and did not change water access decreases from 205 to 176 (86%), from 103 to 80 (78%), and from 2208 to 1627 (74%), respectively. As before, changes in household income are instrumented by changes in rainfall. The estimated coefficient on household income is statistically significant at the ten percent significance level, and implies that the BMI z score increases by 0.03 for each 10% increase in household income. The coefficient for "near water" decreases to 0.159 but is still significant at the ten percent significance level.

Table 3.7 shows the results of the sensitivity analysis to different cutoff criteria. The coefficient estimates on "near water" are presented in the table when we use only communities where no sample households changed water access, or more than 0, 5, 10, 15, 20, and 25 percentages of sample households changed water access in one direction between 1991 and 1993.[12] The estimated effect of access to clean water on child health is U shaped as the cutoff criterion increases. Given the magnitudes of the standard errors,

[11] If the number of sample households within a community differs between the two sample years, the minimum of them is used as the base to calculate the percentage of households that gained/lost "near water" within the community.
[12] The zero-percent cutoff criterion means that all sample households are used.

however, the differences among the estimated effects could be due to random errors. The estimated effects range from 0.16 to 0.20, and are all statistically significant at least at the ten percent significance level no matter which specific cutoff criterion is used.

3.4 Parental Education and Water Access

To examine whether parental education and water access are complements or substitutes, three dummies for maternal education (household head's wife did not graduate from primary school, graduated from primary school, and more than primary school degree) are interacted with dummies for water access. Column (5) reports the results. Household income is instrumented, and I use only communities where no sample households changed water access, or more than fifteen percent of sample households changed water access in one direction between the sample years. The effect of water access on child health is largest for mothers in the highest educational group (lower middle school education or higher degrees). The effect of access to clean water on child health for the most educated group of mothers is a 0.398 increase in the BMI z score and statistically different from zero at the one percent significance level. For the sample households, maternal education and access to clean water appear to be complements.

Column (6) in Table 3.5 interacts water access with paternal education. Three dummies for paternal education (household head did not graduate from primary school, graduated from primary school, and more than primary school degree) are created. Again, the effect of water access on child health is largest for the most educated group of fathers (a 0.259 increase in the BMI z score), although the effect is smaller than the effect of water access for mothers in the highest educational group (0.398). We may be able to consider that paternal education is highly correlated with (time-invariant) permanent household income. Then, the result could imply that the effect of access to clean water is largest for relatively rich households rather than poor households.

3.5 Non-Random Project Placement

The next concern is that even if changes in water access are exogenous to households' own demand, the placement of water projects may not be random. Although we included some dynamic community characteristics that may affect child health in our earlier regression models, we are still concerned that some unobserved changes in community characteristics are correlated with both changes in child health and changes in water access. Note that project placements with respect to time-invariant characteristics do not bias the estimated effect of water accessibility on child health in Equation (1), because child fixed effects fully control for such sources of the bias. The problem arises if local governments allocate water projects in areas where other health-related infrastructure is deteriorating, or if changes in community wealth lead to investments in water projects and other investments (e.g. health facilities) that also affect child health.

To deal with this potential problem, I include community dummies in Equation (1). As long as water project sites are selected using community characteristics (either dynamic or time-invariant characteristics) or characteristics of more aggregated areas, the coefficient estimates are unbiased. In this case, the identification of the effect of water access on child health comes from changes in child health in those households which changed water access versus changes in child health in those households which did not change water access within the same community. For the rest of the chapter, we continue to restrict the sample households to those who lived in communities where no sample households experienced changes in water access, or more than fifteen percent of sample households changed water access in one direction between the two sample years. With the inclusion of community dummies, however, the identification of the effect of water access on child health comes solely from communities where some households experienced changes in water access. Under our identification strategy with community fixed effects, the treatment group consists of children who experienced changes in water access and lived in communities where more than fifteen percent of sample households changed water access, while the control group is groups of children who did not experience changes in water access, and lived in the same communities as counterpart children in the treatment group. This identification strategy is valid even if different communities would have experienced, on average, different changes in child health under the counter factual scenario that no households changed water access between the sample years. As long as the placements of water projects are random within communities after excluding communities where fifteen percent or less sample households experienced changes in water access, our identification is unbiased. A sensitivity analysis to different cutoff criteria with community fixed effects in Appendix 3.1 confirms our earlier finding that the estimated effect of access to clean water on child health is U shaped as the cutoff criterion increases. Again, differences among the estimated effects with the different cutoff criteria could be due to random errors, given the magnitudes of the standard errors.

The downside of including community dummies in Equation (1) is that our rainfall instruments, which are county-level variables, can no longer be used. To address the potential endogeneity of household income, we look at the sample of children of age ten or younger in 1991 in addition to the sample of children of age fifteen or younger in 1991. Also, as long as no omitted variables are correlated with changes in water access, the simultaneous determination of household income and child health does not bias the estimated effect of access to clean water on child health.

With the inclusion of community dummies, Column (1) in Table 3.8 reports that the estimated effect of access to clean water on child health is about 45% smaller than the estimated effect without including community dummies (0.087 versus 0.159) while the standard error is slightly smaller (0.073 versus 0.083). This is evidence that changes in "near water" could be positively correlated with other dynamic characteristics of communities that positively affect child health. This would be the case, for example, if communities that become rich enough to invest in water access also can afford investments in other health-related facilities. However, the difference between the estimated effects (0.087 versus 0.159) is not statistically significant at the conventional significance levels with the p-value equal to 0.32.

Column (2) restricts the sample children to those with age ten or younger in 1991 to avoid the simultaneous determination of household income and child health. The estimated effect of access to clean water on child health for younger children is larger than for children with fifteen or younger (0.120 versus 0.087) although the standard error is also larger (0.092 versus 0.073).

Columns (3) and (4) interact water access with maternal and paternal education, respectively. We confirm that the effect of access to clean water on child health is largest for the most educated group of parents and statistically significant at the five percent significance level for mothers and at the ten percent level for fathers. Including community dummies lowers the magnitudes of the estimated effect of water access on child health for all educational levels except for the least educated group of mothers and the middle educational group of fathers (Column (5) in Table 3.5 versus Column (3) in Table 3.8 for maternal education; and Column (6) in Table 3.5 versus Column (4) in Table 3.8 for paternal education). However, the differences in the estimated effects with and without community fixed effects are not statistically significant at the conventional significance levels for any educational group of fathers and mothers. The results for the sample children with age ten or younger in 1991 are similar (not reported).

Finally, Columns (5) and (6) estimate the effect of water access on child health separately for those who gained and lost access to clean water. Column (5) uses sample children with age fifteen or younger in 1991, and Column (6) is with sample children with age ten or younger. For both children aged fifteen or younger and those aged ten or younger, the effect of water access on child health is not symmetric in terms of gained and lost access, although the differences in the estimated coefficients are not statistically significant (The p-values for the tests of the differences in the estimated coefficients are 0.58 for children with age fifteen or younger and 0.93 for children with age ten or younger). It could be that losing water access impairs child health more than gaining access increases it.

Our results suggest that unobserved community dynamics positively affecting child health could be positively correlated with changes in water access, so ignoring such community dynamics could bias upward the estimated effect of water access on child health. We directly check whether some observable community dynamics (but not included in the regressions) are correlated with changes in water access at the community level. Table 3.9 presents simple correlations between changes in the percentage of households with "near water" and changes in various community characteristics. For each community characteristic, two correlations are shown. The first correlation is calculated using all sample communities, while the second correlation is calculated using the sample communities where no sample households changed water access, or more than fifteen percent of households changed water access in one direction between 1991 and 1993. Generally, changes in water access are rarely correlated with changes in community characteristics to the extent that the correlations are statistically significant. However, if we focus on the few observed correlations that are statistically significant, they seem to support our finding.

3.6 Our Identification Strategies Effective?

Finally, we would like to have a sense of what types of children gained or lost water access between 1991 and 1993 and under what conditions our identification strategies are valid. Our identification strategies attempt to choose the treatment and control groups of children, so that they would have experienced, on average, the same trend in the BMI z score under the counter factual scenario that no children changed water access between 1991 and 1993. Our treatment group consists of children who gained or lost water access between the sample years, and our control group consists of sample children who experienced no changes in water access. Table 3.10 presents the results of the probit regressions where the dependent variable is one if the household belongs to the treatment group and zero otherwise. The right-hand-side variables contain both household and community characteristics including time-invariant characteristics, initial characteristics in 1991, and dynamic characteristics (changes). In parallel with our identification strategies, Column (1) uses all communities without community fixed effects, Column (2) uses chosen communities with the fifteen percent restriction without community fixed effects, and Column (3) uses chosen communities with the fifteen percent restriction with community fixed effects. Of course, we can use only observable characteristics in the probit regressions, and we remain ignorant about whether unobserved household and community characteristics are systematically different between the treatment and control groups.

Column (1) shows that the treatment and control groups are systematically different in some household and community characteristics. Column (1) implies that our first identification strategy (with all communities and without community fixed effects) requires for unbiasedness that children experienced the same trend in the BMI z score on average regardless of the levels of parental education and some initial and dynamic community characteristics in which they live.[13] Column (2) shows that the treatment and

[13] These conditions are not sufficient but only necessary in the sense that there are many unobserved characteristics that are potentially correlated with changes in child health and changes in water access.

control groups could be more systematically different in terms of observed characteristics under our second identification strategy (with chosen communities and without community fixed effects). Column (3) includes community fixed effects and controls for all inter-community differences. Still, we see some systematic differences in household characteristics between the treatment and control groups. Particularly, our third identification strategy (with chosen communities and with community fixed effects) needs for unbiasedness that sample children would have experienced the same counter-factual BMI trend even when parental education and the initial BMI z score are different.

4. Conclusions

The effect of water accessibility on child health is measured using child-specific fixed effects in this chapter. Fixed effects control for unobserved time-invariant characteristics that could bias the estimated effect of water access on child health. In addition, we address dynamic characteristics of both households and communities that could bias the results. To control for unobserved household dynamics that could be correlated with both changes in water access and changes in child health, we use a subset of the sample children whose water access was likely affected by external projects rather than households' own demand. To deal with unobserved community dynamics that could affect child health, we make use of community dummies.

We do not have strong evidence about the directions of biases in case we neglect unobserved household and community dynamics that are potentially correlated with both changes in child health and changes in water access. Our results show that although unobserved household dynamics push the estimated effect of water access on child health upward, the difference is not statistically significant at the conventional significance levels. Addressing unobserved community dynamics also decreases the estimated effects of water access on child health, but the difference is not statistically significant, either.

After controlling for confounding household and community dynamics, the magnitude of the estimated effect becomes smaller to 0.087 or 0.120, which implies that having access to clean water within the yard of one's house increases the child BMI z score by 0.087 or 0.120 (Columns (1) and (2) in Table 3.8). To obtain sounder statistical ground for the estimates, we need a larger sample size. This study also finds that losing water access could impair child health more than gaining access increases it. It would be important not only to provide access to clean water within the yard of one's house but also to maintain the source of clean water people already gained within the yards of their houses.

To verify our identification strategies, we look at differences between the treatment and control groups using the probit regressions. Our results find that even with the most preferred identification strategy (with chosen communities and with community fixed effects), differing child and household characteristics between the treatment and control groups could bias the results. Particularly, our identification strategy requires for unbiasedness that sample children would have experienced the same counter-factual BMI trend even when parental education and the initial BMI z score are different.

We find no consistent estimate of the effect of changes in household income on changes in child health. Using measures of rainfall variation as instruments for household income, we have the expected positive and statistically significant effect of household income on child health. However, we are not confident about why the base-line child fixed effects model biases downward the effect of household income on child health. Inter-temporal consumption smoothing may dampen the effect of changes in household income on changes in child health.

We also find that "near water" and parental education would be complements in producing child health for the sample children. In other words, the effect of "near water" on child health is largest for children with most educated parents.

A straightforward policy implication of this study is that providing access to clean water within the yard of one's house and maintaining it would increase child health. Also, increasing parental schooling would make water projects more effective to improve child health.

Besides examining whether convenient access to clean drinking water improves child health, how it affects child health is also policy relevant. One plausible mechanism through which access to clean water affects child health is the improved quality of water and thus the reduced contact to germs. If the source of water is physically proximate to dwellings, people do not need to store water, and fresh water is always available. Alternatively, water sources within the yard of one's house (typically, wells and processed water from the tap) may be better in quality in the first place than water obtained outside of one's house (such as rivers, ponds, and public wells). It is also possible that the quantity as well as the quality of water matter for child health. Convenient access to clean water would help people obtain the amount of water they need whenever they need. Finally, convenient access to clean water may release mothers from fetching water. If they use the saved time for child care or more generally home production, child health may increase as a result. As a direction of further studies, exploring how water access affects health would be useful in better policy making.

Table 3.1: Description of Variables

1. BMI z score (a)	BMI standardized by age and sex using healthy children in the US as the reference population. Children whose BMI z scores are 3 or greater in absolute value in at least one sample year are not used due to highly likely misreports of heights and/or weights. Also, children whose BMI z scores are 2 or greater in at least one sample year are not used to exclude obese children.
2. male (a)	Equals 0 if the child is female and equals 1 if the child is male.
3. child age (a)	Child age in years ranging from 2 to 15 in 1991.
4. near water (b)	Equals 1 if the household has in-house tap water, in-yard tap water, or in-yard well to obtain drinking water; Equals 0 if the household obtains drinking water from other place.
5. log per-capita hh deflated income (b)	Log of per-capita household income deflated by price index provided in the CHNS.
6. village (c)	Equals 1 if the community is categorized as village; Equals 0 if the community is categorized as either urban, suburban, or town.
7. immunization (c)	Ratio of children (age 12 or younger) who received any immunization in the past 12 months to total number of children within the community.
8. preventive health service (c)	Ratio of individuals who received any preventive health service in the past month to total number of individuals within the community.
9. medicine 20 (c)	Ratio of households that say that needed medicine is generally available in a medical facility reachable within 20 minutes, to total number of households within the community.

1) (a) child-level data, (b) household-level data, (c) community-level data

Table 3.2: Summary Statistics of Child/Household/Community Characteristics

	cross-sectional mean for 1991	standard deviation	standard deviation of within-child variation
BMI z score (a)	-0.470	1.009	0.988
child age (a)	8.378	3.869	0.264
log of per-capita hh deflated income (b)	6.586	0.947	1.427
near water (b)	0.810	0.392	0.348
immunization (c)	0.606	0.236	0.284
preventive health service (c)	0.019	0.039	0.050
medicine 20 (c)	0.837	0.228	0.200

time-invariant characteristics:	% of total
gender (a)	
male	52.4
female	47.6
educational level of hh head (b)	
less than primary school	14.0
graduated from primary school	22.1
lower middle school or higher	63.9
educational level of hh head's wife (b)	
less than primary school	35.8
graduated from primary school	20.6
lower middle school or higher	43.7
residence (c)	
village	61.1
urban, suburban, or town	38.9
province	
Liaoning	11.2
Henan	10.2
Shandong	9.7
Hubei	14.1
Hunan	13.4
Jiangsu	9.7
Guangxi	15.7
Guizhou	15.9
1991 interview conducted in (b)	
August	1.0
September	38.6
October	43.3
November	16.8
December	0.4
1993 interview conducted in (b)	
June	0.1
August	1.2
September	32.7
October	38.9
November	22.2
December	5.0

1) (a) child-level data, (b) household-level data, (c) community-level data
2) The number of the sample children used to calculate the statistics is 2474 ~ 2521.
3) If parental education differs between 1991 and 1993, the max of them is reported.

Table 3.3: Sample Children by "Near Water" Status in 1991 and 1993

near water in 91 => near water in 93	# children	(%)
Yes => Yes	1936	(76.95)
Yes => No	103	(4.09)
No => Yes	205	(8.15)
No => No	272	(10.81)
Total	2516	(100.00)

Table 3.4: Cross-Tabulation of Changes in Water Access and Changes in BMI z Score

near water in 91 => near water in 93	Δ z<0		Δ z>=0		Total	
Yes => Yes	1001	(51.70%)	935	(48.30%)	1936	(100.00%)
Yes => No	50	(48.54%)	53	(51.46%)	103	(100.00%)
No => Yes	82	(40.00%)	123	(60.00%)	205	(100.00%)
No => No	152	(55.88%)	120	(44.12%)	272	(100.00%)
Total	1285	(51.07%)	1231	(48.93%)	2516	(100.00%)

Table 3.5: Econometric Results

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	15
Income instrumented?	No	No	Yes	Yes	Yes	Yes
Interacted with Maternal or Paternal Education?	No	No	No	No	Yes	Yes
All or Chosen communities?	All	All	All	Chosen	Chosen	Chosen
Δ water access within yard (Δ near water)	0.178***	0.152*	0.204***	0.159*		
	(0.066)	(0.085)	(0.078)	(0.083)		
(Δ near water) X (hh head's wife, less than primary school education)					0.006	
					(0.111)	
(Δ near water) X (hh head's wife, primary school education)					0.189	
					(0.211)	
(Δ near water) X (hh head's wife, lower middle school education or more)					0.398***	
					(0.142)	
(Δ near water) X (hh head, less than primary school education)						0.073
						(0.192)
(Δ near water) X (hh head, primary school education)						0.014
						(0.169)
(Δ near water) X (hh head, lower middle school education or more)						0.259***
						(0.097)
Δ log (deflated hh per-capita income)	-0.003	-0.015	0.347***	0.319*	0.328**	0.249
	(0.018)	(0.022)	(0.125)	(0.166)	(0.164)	(0.152)
Δ immunization	-0.178**	-0.158	-0.262***	-0.299***	-0.309***	-0.292***
	(0.078)	(0.106)	(0.097)	(0.111)	(0.109)	(0.106)
Δ preventive health service	0.521	1.080*	0.118	-0.049	-0.117	0.083
	(0.509)	(0.622)	(0.583)	(0.666)	(0.665)	(0.650)
Δ medicine 20	0.083	0.084	-0.002	-0.053	-0.062	-0.026
	(0.118)	(0.159)	(0.137)	(0.159)	(0.158)	(0.152)

Table 3.5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	15
Income instrumented?	No	No	Yes	Yes	Yes	Yes
Interacted with Maternal or Paternal Education?	No	No	No	No	Yes	Yes
All or Chosen communities?	All	All	All	Chosen	Chosen	Chosen
2-3 years old in 1991 (dummy)	-0.180**	-0.176*	-0.162*	-0.167*	-0.165*	-0.192**
	(0.074)	(0.096)	(0.085)	(0.092)	(0.092)	(0.087)
4-6 years old in 1991 (dummy)	-0.275***	-0.273***	-0.265***	-0.301***	-0.304***	-0.302***
	(0.060)	(0.086)	(0.067)	(0.076)	(0.077)	(0.075)
7-9 years old in 1991 (dummy)	-0.080	-0.079	-0.065	-0.068	-0.065	-0.049
	(0.054)	(0.082)	(0.062)	(0.070)	(0.071)	(0.068)
10-12 years old in 1991 (dummy)	0.041		0.056	0.065	0.061	0.061
	(0.052)		(0.056)	(0.064)	(0.065)	(0.063)
Male (dummy)	-0.064	-0.034	-0.084*	-0.081*	-0.074	-0.074
	(0.039)	(0.051)	(0.044)	(0.049)	(0.050)	(0.048)
Village (dummy)	0.020	0.033	0.095*	0.132*	0.128*	0.122*
	(0.045)	(0.060)	(0.057)	(0.068)	(0.068)	(0.064)
Liaoning (dummy)	0.271***	0.317**	0.398***	0.257*	0.257*	0.255*
	(0.097)	(0.124)	(0.124)	(0.135)	(0.135)	(0.132)
Henan (dummy)	-0.085	-0.108	0.232	0.113	0.124	0.070
	(0.089)	(0.118)	(0.164)	(0.184)	(0.183)	(0.175)
Shandong (dummy)	-0.015	-0.078	0.234	0.160	0.160	0.091
	(0.094)	(0.128)	(0.143)	(0.178)	(0.176)	(0.175)
Hubei (dummy)	0.174**	0.088	0.251**	0.107	0.109	0.111
	(0.087)	(0.122)	(0.103)	(0.110)	(0.111)	(0.106)
Hunan (dummy)	0.146	0.210	0.262**	0.265*	0.257*	0.285**
	(0.099)	(0.141)	(0.123)	(0.140)	(0.141)	(0.135)

Table 3.5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	15
Income instrumented?	No	No	Yes	Yes	Yes	Yes
Interacted with Maternal or Paternal Education?	No	No	No	No	Yes	Yes
All or Chosen communities?	All	All	All	Chosen	Chosen	Chosen
Guangxi (dummy)	0.137	0.078	0.141	0.018	0.027	0.014
	(0.088)	(0.124)	(0.101)	(0.112)	(0.114)	(0.107)
Guizhou (dummy)	0.073	0.073	0.021	0.019	0.024	0.033
	(0.077)	(0.112)	(0.088)	(0.095)	(0.095)	(0.091)
Interviewed in September or before	0.220***	0.282***	0.345***	0.308***	0.318***	0.295***
in 1991 (dummy)	(0.055)	(0.073)	(0.079)	(0.090)	(0.090)	(0.084)
Interviewed in November or December	-0.087	-0.029	-0.054	-0.156	-0.144	-0.114
in 1991 (dummy)	(0.073)	(0.100)	(0.079)	(0.104)	(0.104)	(0.102)
Interviewed in September or before	-0.338***	-0.423***	-0.363***	-0.339***	-0.339***	-0.328***
in 1993 (dummy)	(0.061)	(0.083)	(0.072)	(0.078)	(0.078)	(0.074)
Interviewed in November or December	-0.046	-0.028	-0.158**	0.012	0.002	0.004
in 1993 (dummy)	(0.060)	(0.085)	(0.079)	(0.080)	(0.080)	(0.077)
Constant	0.050	0.033	-0.047	0.003	-0.004	-0.002
	(0.085)	(0.128)	(0.105)	(0.117)	(0.117)	(0.110)
Observations	2516	1662	2516	1883	1874	1842
F statistic on the excluded instruments			F(2,1713)=12.86	F(2,1312)=7.72	F(2,1305)=7.78	F(2,1281)=8.11
p-value Hansen J statistic (over-id test statistic)			χ^2 (1)P-val =0.32	χ^2 (1)P-val =0.96	χ^2 (1)P-val =0.95	χ^2 (1)P-val =0.62

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.

2) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.5 (continued)

Joint significance of excluded IV's in regressing changes in farm work hours on the first-stage exogenous variables

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	15
Income instrumented?	No	No	Yes	Yes	Yes	Yes
Interacted with Maternal or Paternal Education?	No	No	No	No	Yes	Yes
All or Chosen communities?	All	All	All	Chosen	Chosen	Chosen
F statistic on the excluded instruments			F(2,1710)=0.65	F(2,1309)=0.07	F(2,1302)=0.06	F(2,1278)=0.08
p-value joint significance of the excluded IV's			0.52	0.93	0.94	0.93
Observations			2504	1874	1865	1833

Table 3.6: Number of Sample Households within Communities

# households reporting water access	1991 # communities	%	1993 # communities	%
9			1	(0.56%)
10				
11			3	(1.68%)
12			1	(0.56%)
13	1	(0.53%)	2	(1.12%)
14	1	(0.53%)	3	(1.68%)
15	5	(2.63%)	6	(3.35%)
16	3	(1.58%)	13	(7.26%)
17	13	(6.84%)	12	(6.70%)
18	29	(15.26%)	28	(15.64%)
19	38	(20.00%)	40	(22.35%)
20	100	(52.63%)	60	(33.52%)
21			4	(2.23%)
22			5	(2.79%)
23			1	(0.56%)
Total	190	(100%)	179	(100%)

Table 3.7: Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria

Cutoff	Coeff. Sig	Robust SE	N Total	%	N Gain Access	%	N Lost Access	%	N No Change in Access	%
0%	0.204***	0.078	2516	100.0%	205	8.1%	103	4.1%	2208	87.8%
5%	0.190**	0.078	2380	100.0%	204	8.6%	101	4.2%	2075	87.2%
10%	0.163**	0.076	2100	100.0%	194	9.2%	96	4.6%	1810	86.2%
15%	0.159*	0.083	1883	100.0%	176	9.4%	80	4.3%	1627	86.4%
20%	0.186**	0.093	1704	100.0%	160	9.4%	73	4.3%	1471	86.3%
25%	0.204*	0.106	1597	100.0%	146	9.1%	55	3.4%	1396	87.4%

1) * significant at the ten percent; ** significant at the five percent; *** significant at the one percent

2) Change in household income (one of the covariates) is instrumented using county-level rainfall variation.

Table 3.8: Econometric Results with Community Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	10
Income instrumented?	No	No	No	No	No	No
Interacted with Parental Education?	No	No	Yes	Yes	No	No
All or Chosen communities?	Chosen	Chosen	Chosen	Chosen	Chosen	Chosen
Δ water access within yard (Δ near water)	0.087	0.120				
	(0.073)	(0.092)				
(Δ near water) X (hh head's wife, less than primary school education)			0.014			
			(0.094)			
(Δ near water) X (hh head's wife, primary school education)			0.058			
			(0.163)			
(Δ near water) X (hh head's wife, lower middle school education or more)			0.241**			
			(0.111)			
(Δ near water) X (hh head, less than primary school education)				-0.051		
				(0.139)		
(Δ near water) X (hh head, primary school education)				0.035		
				(0.143)		
(Δ near water) X (hh head, lower middle school education or more)				0.159*		
				(0.083)		
(Δ near water = gain)					0.049	0.112
					(0.099)	(0.124)
(Δ near water = lost)					0.145	0.132
					(0.126)	(0.161)

Table 3.8 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	10
Income instrumented?	No	No	No	No	No	No
Interacted with Parental Education?	No	No	Yes	Yes	No	No
All or Chosen communities?	Chosen	Chosen	Chosen	Chosen	Chosen	Chosen
Δ log (deflated hh per-capita income)	0.002	-0.008	0.000	0.003	0.002	-0.008
	(0.016)	(0.018)	(0.016)	(0.016)	(0.016)	(0.018)
2-3 years old in 1991 (dummy)	-0.255***	-0.344***	-0.251***	-0.277***	-0.252***	-0.344***
	(0.088)	(0.115)	(0.089)	(0.089)	(0.088)	(0.115)
4-6 years old in 1991 (dummy)	-0.298***	-0.389***	-0.302***	-0.294***	-0.298***	-0.389***
	(0.067)	(0.099)	(0.067)	(0.068)	(0.067)	(0.099)
7-9 years old in 1991 (dummy)	-0.104	-0.186*	-0.102	-0.088	-0.103	-0.186*
	(0.065)	(0.099)	(0.066)	(0.066)	(0.065)	(0.099)
10-12 years old in 1991 (dummy)	0.094		0.091	0.093	0.094	0.000
	(0.062)		(0.062)	(0.062)	(0.062)	0.000
Male (dummy)	-0.051	0.017	-0.043	-0.040	-0.050	0.018
	(0.044)	(0.058)	(0.045)	(0.044)	(0.044)	(0.058)
Interviewed in September or before	-0.013	0.055	0.000	0.000	-0.008	0.057
in 1991 (dummy)	(0.200)	(0.173)	(0.200)	(0.201)	(0.202)	(0.175)
Interviewed in November or December	-0.169	0.011	-0.155	-0.158	-0.155	0.015
in 1991 (dummy)	(0.207)	(0.269)	(0.206)	(0.208)	(0.206)	(0.276)
Interviewed in September or before	-0.830***	-1.695***	-0.810***	-0.810***	-0.845***	-1.696***
in 1993 (dummy)	(0.275)	(0.373)	(0.271)	(0.266)	(0.279)	(0.373)
Interviewed in November or December	-0.314	0.009	-0.323	-0.336	-0.309	0.009
in 1993 (dummy)	(0.425)	(0.465)	(0.426)	(0.427)	(0.426)	(0.465)

Table 3.8 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Max age of children in 1991?	15	10	15	15	15	10
Income instrumented?	No	No	No	No	No	No
Interacted with Parental Education?	No	No	Yes	Yes	No	No
All or Chosen communities?	Chosen	Chosen	Chosen	Chosen	Chosen	Chosen
Constant	0.383	1.868***	0.369	0.369	0.369	1.863***
	(0.392)	(0.570)	(0.392)	(0.393)	(0.392)	(0.574)
Observations	1883	1234	1874	1842	1883	1234

1) Robust standard errors in parentheses are robust to household-level clustering and heteroskedasticity.

2) All columns include community dummies (the coefficient estimates not reported) besides the variables listed above.

3) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.9: Community-level Dynamic Correlation between Changes in % Households with Access to "Near Water" and Changes in Various Community Characteristics

	Correlation	Sig.	p-value	N	
Δ % hh's that say that needed medicine generally available	0.157	**	0.032	188	Included in the regressions
in a medical facility reachable within 20 minutes	0.183	**	0.026	149	
Δ % children (age 12 or younger) who received	-0.068		0.357	188	
any immunizations in the past 12 months	-0.061		0.464	149	
Δ % individuals who received any preventive	-0.015		0.843	188	
health service in the past month	-0.009		0.913	149	
Δ log (mean per-capita hh deflated income)	-0.060		0.415	188	
	-0.081		0.324	149	
Δ number patient beds	-0.078		0.310	173	
	-0.085		0.327	135	
Δ time cost to the nearest medical facility (minutes;	-0.001		0.893	188	
hh mean)	-0.027		0.740	149	
Δ most common characteristics of road being paved	-0.043		0.579	168	
road (dummy)	-0.049		0.582	129	
Δ convenient telephone service available (dummy)	0.080		0.286	181	
	0.078		0.355	142	
Δ daily provincial newspaper available on the day it	0.164	**	0.027	181	
is published (dummy)	0.161	*	0.056	142	
Δ convenient to see movies (dummy)	0.150	**	0.045	180	
	0.165	*	0.051	141	
Δ how many hours per day electricity available	0.091		0.225	179	
on average	0.120		0.157	140	
Δ how many days a week electricity cutoff	-0.098		0.198	176	
on average	-0.118		0.167	138	
Δ % workforce engaging mainly in agricultural	-0.173	**	0.023	173	
activity	-0.177	**	0.041	134	
Δ % workforce who worked out of town for	0.132	*	0.085	171	
more than 1 month last year	0.137		0.119	132	
Δ % workforce who work in enterprises employing	-0.064		0.416	162	
more than 20 people	-0.082		0.364	125	
Δ % workforce who work in enterprises employing	-0.010		0.903	161	
fewer than 20 people	0.002		0.980	126	
Δ number village/town/county/neighborhood	-0.008		0.917	178	
enterprises	0.003		0.973	139	
Δ number self-employed household enterprises	0.057		0.453	175	
	0.061		0.481	138	
Δ number private enterprises	-0.011		0.890	175	
	-0.009		0.917	136	
Δ number in-door restaurants	0.002		0.978	178	
	-0.002		0.983	140	

1) For each community characteristic, two simple correlations are shown. The upper row uses all sample communities, while the lower row uses the sample communities where no sample households changed water access, or more than 15% of households changed water access in one direction between 1991 and 1993.

2) * significant at the 10%; ** significant at the 5%; *** significant at the 1%

Table 3.10: Probit Regressions of Treatment (Changes in Water Access)

		(1)	(2)	(3)
	All or Chosen communities?	All	Chosen	Chosen
	Community fixed effects included?	No	No	Yes
household characteristics	mean child BMI z score in 1991	0.012	0.022	0.227**
		(0.058)	(0.073)	(0.114)
	ratio of boys to total number of children	0.061	0.203	0.221
		(0.118)	(0.154)	(0.211)
	mean child age in 1991	-0.017	-0.044**	-0.037
		(0.015)	(0.019)	(0.027)
	log per-capita deflated income in 1991	0.037	-0.015	-0.008
		(0.062)	(0.078)	(0.131)
	Δ log per-capita deflated income	-0.001	-0.012	0.086
		(0.043)	(0.053)	(0.065)
	hh head, primary school education (dummy)	0.428**	0.594**	0.595**
		(0.182)	(0.231)	(0.236)
	hh head, lower middle school education or more (dummy)	0.255	0.178	0.211
		(0.180)	(0.228)	(0.253)
	hh head's wife, primary school education (dummy)	-0.273*	-0.260	-0.380*
		(0.144)	(0.181)	(0.228)
	hh head's wife, lower middle school education or more (dummy)	-0.224	-0.334*	-0.427*
		(0.141)	(0.173)	(0.240)
community characteristics	village resident (dummy)	0.244	0.309	
		(0.159)	(0.208)	
	Liaoning (dummy)	-0.211	0.155	
		(0.211)	(0.262)	
	Shandong (dummy)	-0.797**	-1.431***	
		(0.331)	(0.501)	
	Henan (dummy)	-0.333	-0.649**	
		(0.237)	(0.297)	
	Hubei (dummy)	0.364*	0.606**	
		(0.202)	(0.281)	
	Hunan (dummy)	-0.114	-0.133	
		(0.204)	(0.282)	
	Guangxi (dummy)	0.041	0.266	
		(0.215)	(0.329)	
	Guizhou (dummy)	-0.255	-0.752**	
		(0.261)	(0.366)	
	percentage workforce engaged in agriculture in 1991	0.001	0.000	
		(0.002)	(0.003)	
	Δ percentage workforce engaged in agriculture	0.002	0.000	
		(0.002)	(0.003)	
	log mean per-capita deflated hh income in 1991	-0.180	-0.252	
		(0.162)	(0.216)	
	Δ log mean per-capita deflated hh income	-0.337**	-0.639***	
		(0.143)	(0.204)	

Table 3.10 (continued)

		(1)	(2)	(3)
	All or Chosen communities?	All	Chosen	Chosen
	Community fixed effects included?	No	No	Yes
community characteristics	(number of patient beds in 1991) /1000	-0.472	-0.293	
		(0.308)	(0.401)	
	Δ number of patient beds/1000	-1.250***	-1.002**	
		(0.363)	(0.397)	
	ratio of households with near water in 1991	-1.078***	-2.150***	
		(0.200)	(0.300)	
	ratio of children (age 12 or younger) who received any immunization in the past 12 months in 1991	-0.001	-0.218	
		(0.319)	(0.414)	
	Δ ratio of children (age 12 or younger) who received any immunization in the past 12 months	-0.197	-0.770***	
		(0.244)	(0.279)	
	ratio of individuals who received any preventive health service in the past month in 1991	-0.607	-2.515	
		(2.445)	(3.180)	
	Δ ratio of individuals who received any preventive health service in the past month	-3.512	-5.168*	
		(2.231)	(2.785)	
	ratio of hh's that say that needed medicine is generally available in a medical facility reachable within 20 minutes in 1991	-0.162	-1.999***	
		(0.546)	(0.764)	
	Δ ratio of hh's that say that needed medicine is generally available in a medical facility reachable within 20 minutes	1.090*	-1.678*	
		(0.579)	(0.880)	
	mean time cost to the closest medical facility in 1991	-0.003	-0.023	
		(0.012)	(0.018)	
	Δ mean time cost to the closest medical facility	0.009	-0.034**	
		(0.011)	(0.013)	
	Constant	0.576	4.389***	-0.691
		(1.188)	(1.660)	(0.972)
	Observations	1403	1051	326

1) Standard errors in parentheses are robust to heteroskedasticity.

2) Column (3) includes community dummies (the coefficient estimates not reported) besides the variables listed above.

3) * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix 3.1 Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects

Table 3.11: Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects (Children Age 15 or Younger)

Cutoff	Coeff. Sig	Robust SE	Total N	Total %	Gain Access N	Gain Access %	Lost Access N	Lost Access %	No Change in Access N	No Change in Access %
0%	0.103	0.068	2516	100.0%	205	8.1%	103	4.1%	2208	87.8%
5%	0.090	0.068	2380	100.0%	204	8.6%	101	4.2%	2075	87.2%
10%	0.096	0.071	2100	100.0%	194	9.2%	96	4.6%	1810	86.2%
15%	0.087	0.073	1883	100.0%	176	9.4%	80	4.3%	1627	86.4%
20%	0.101	0.079	1704	100.0%	160	9.4%	73	4.3%	1471	86.3%
25%	0.102	0.083	1597	100.0%	146	9.1%	55	3.4%	1396	87.4%

1) * significant at the ten percent; ** significant at the five percent; *** significant at the one percent

Table 3.12: Sensitivity Analysis of the Estimated Effect of "Near Water" to Differing Cutoff Criteria with Community Fixed Effects (Children Age 10 or Younger)

Cutoff	Coeff. Sig	Robust SE	Total N	Total %	Gain Access N	Gain Access %	Lost Access N	Lost Access %	No Change in Access N	No Change in Access %
0%	0.153*	0.089	1662	100.0%	126	7.6%	76	4.6%	1460	87.8%
5%	0.147	0.090	1584	100.0%	126	8.0%	74	4.7%	1384	87.4%
10%	0.154*	0.092	1389	100.0%	120	8.6%	72	5.2%	1197	86.2%
15%	0.120	0.092	1234	100.0%	108	8.8%	60	4.9%	1066	86.4%
20%	0.129	0.101	1107	100.0%	98	8.9%	54	4.9%	955	86.3%
25%	0.151	0.102	1038	100.0%	88	8.5%	40	3.9%	910	87.7%

1) * significant at the ten percent; ** significant at the five percent; *** significant at the one percent

References

Behrman, Jere, and Barbara Wolfe. (1987). "How Does Mother's Schooling affect Family Health, Nutrition, Medical Care Usage, and Household Sanitation?" *Journal of Econometrics*, 36: 185-204.

Behrman, Jere, and Barbara Wolfe. (1989). "Does More Schooling Make Woman Better Nourished and Healthier? Adult Sibling Random and Fixed Effects Estimates for Nicaragua," *Journal of Human Resources*, 24 (4): 644-663.

Cebu Study Team. (1991). "Underlying and Proximate Determinants of Child Health: the Cebu Longitudinal Health and Nutrition Study," *American Journal of Epidemiology*, 133 (2): 185-201.

China Daily. (2004). "Goal: Clean Drinking Water for All by 2020," November 29, 2004.

Datcher-Loury, Linda. (1988). "Effects of Mother's Home Time on Children's Schooling," *Review of Economics and Statistics*, 70 (3): 367-373.

De Zuane, J. (1997). *Handbook of Drinking Water Quality*, New York: Van Nostrand.

Fafchamps, Marcel. (1993). "Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers," *Econometrica*, 61 (5): 1173-97.

Frankenberg, Elizabeth, and Duncan Thomas. (2001). "Women's Health and Pregnancy Outcomes: Do Services Make a Difference?" *Demography*, 38 (2): 253-265

Gadgil, Ashok. (1998). "Drinking Water in Developing Countries," *Annual Review of Energy and the Environment*, 23: 253-86.

Jalan, Jyotsna, and Martin Ravallion. (2001). "Does Piped Water Reduce Diarrhea for Children in Rural India?" mimeo, Development Research Group, The World Bank.

Leibowitz, Arleen. (1974). "Education and Home Production," *American Economic Review*, 64 (2): 243-250.

Lee, Lung-fei, Mark Rosenzweig, and Mark Pitt. (1997). "The Effects of Improved Nutrition, Sanitation, and Water Quality on Child Health in High-Mortality Populations," *Journal of Econometrics*, 77: 209-235.

Merrick, Thomas. (1985). "The Effect of Piped Water on Early Childhood Mortality in Urban Brazil, 1970 to 1976," *Demography*, 22 (1): 1-24.

Molyneaux, John W., and Paul J. Gertler. (2000). "The Impact of Targeted Family Planning Programs in Indonesia," *Population and Development Review*, 26: 61-85.

Pitt, Mark M., Shahidur R. Khandker, Signe-Mary McKernan, and M. Abdul Latif. (1999). "Credit Programs for the Poor and Reproductive Behavior in Low-Income Countries: Are the Reported Causal Relationships the Result of Heterogeneity Bias?" *Demography*, 36 (1): 1-21.

Skoufias, Emmanuel. (1993). "Seasonal Labor Utilization in Agriculture: Theory and Evidence from Agrarian Households in India," *American Journal of Agricultural Economics*, 75 (1): 20-32.

Thomas, Duncan, John Strauss, and Maria-Helena Henriques. (1991). "How Does Mother's Education Affect Child Height?" *Journal of Human Resources*, 26 (2): 183-211.

Thomas, Duncan, and John Strauss. (1992). "Prices, Infrastructure, Household Characteristics and Child Height," *Journal of Development Economics*, 39: 301-331.

World Health Organization. (1996). *Water Supply and Sanitation Sector Monitoring Report, WHO Report, WHO/EOS/96.*15, Geneva, Switzerland.

World Health Organization, and United Nations Children's Fund. (2000). *Global Water Supply and Sanitation Assessment 2000 Report*, Geneva / New York.